

Are the philistines in charge?

A curious argument in the London Times suggests that those who would make the British universities more productive do not understand what science is about.

SIR Douglas Hague, the chairman of the British research council charged with supporting social science, has been a valuable influence among the grander councils of the research establishment in the past two years. He has rightly raised awkward questions (still not answered) about the management of the British government's support for civil science, in particular asking whether the Advisory Board for the Research Councils is constituted in a way that will allow it to do the job intended for it. Sir Douglas is also a member of the committee looking into the future of British membership of the European centre for high-energy physics at Geneva, and is widely expected to be found among those who advocate pulling out when the report appears in a few weeks. Earlier this month, he argued (in *The Times*) the case for stronger links between industry and academic science, and in passing made a side-swipe at the academics of the University of Oxford for having denied the Prime Minister, Mrs Margaret Thatcher, an honorary degree, which is also fair (see *Nature* 7 February). If Sir Douglas Hague is a philistine, he is both measured and moderate in that role, more a gadfly than a real philistine.

The same cannot unfortunately be said for his supporters, especially Dr K.A. Harrap, who described himself (*The Times*, 19 April) as having been, until recently "head of marketing" at the Natural Environment Research Council. His premise is that British science, which is "fit, able and creative", has for too long been dependent on only the government for support. Dr Harrap continues by acknowledging that there has now been a change of circumstances, but that it will not suffice that the scientific community should simply cry "foul". So, he concludes, scientists must "publicly emphasise their wealth-generating capability" not only in the public interest but so as to "demonstrate to the public and its government that science is good value". For good measure, he asks that "publicly funded scientists themselves" must demonstrate that they can turn "intellectual strength" into wealth, and that if, in the process, the "cutting-edge of blue-skies research becomes somewhat blunted... so be it". The meaning of the phrase is probably clear even if its imagery is a little confusing.

The reasons why this argument is false, and dangerously so, deserve to be widely understood, for there is every reason to suppose that the same mistaken conclusions account for much of the British government's indifference to the damage it has been doing to the scientific enterprise during the past five years. (Dr Harrap, for that reason, should not think he has been unfairly singled out for criticism; indeed, he might even claim credit for having made public, perhaps unwittingly, a line of argument that should be better known.) The issue is the balance that should be struck between basic and applied research in a country such as Britain which has, for far too long, supposed that the economic problems that have afflicted it do not exist, and the question of the role of basic research when wealth-creation, in Dr Harrap's language, is the goal.

No harm is done by conceding Dr Harrap's premise, and its antecedents. British research support has been unduly monolithic. British academics have been too uniformly inclined to measure their success in exclusively academic terms. It might be added that much of the reason why British academic research has been industrially disappointing comes from the unwillingness of

government departments and industrial companies in Britain to credit British academics with being capable of generating valuable innovations. If it were otherwise, the defence and industry departments would be more than the minuscule sources of research support they are at present, while the handful of British universities that have been glad to welcome industrial support would have prospered to the envy of the others. It serves no purpose for anybody to lay Britain's economic poverty at the doors of the academics.

It is also possible to concede the continuation of the argument, that crying foul on a change of circumstances serves no purpose. The British academic community as a whole, not just the scientific part of it, was far too slow to recognize the threats to its integrity when these first began to appear more than a decade ago. Not surprisingly, the pretence that there was nothing wrong has now been overtaken by a sense that everything has turned to dust and ashes, and there is nothing to be done. In its own interests, the academic community in Britain should accept the Hague/Harrap argument that the academic community should recognize the serious character of the government's economic problems, show willing in helping solve them and be more imaginative in exerting itself in its own interests.

So how, when so much is conceded, can the Harrap line be mistaken? Because it makes the false assumption that prescriptions that can fairly be held to apply to the scientific community should apply to all scientists individually. And because it goes on to conclude that it would be possible for a country such as Britain, seeking to make its way in a more competitive world by harnessing all "publicly funded" researchers to the same goal of wealth creation, to abandon imaginative research with no foreseeable application without quickly running into trouble. Both conclusions are false. Present circumstances probably require (and will certainly drive) more academics to look for industrial projects on which to work, but why should the same recipe apply to each and all of them? And the second, that imaginative research is "for the time being" dispensable, is untenable, at least so long as it is intended that the education of students should be more than mere paper and that the applied research of future decades has some grist to feed its mill.

This is where the philistines go wrong. They forget, or they choose to overlook, that the scientific enterprise is all of a piece, and that while the balance between the different parts of it can be changed, the belief that some parts of it can be dispensed with altogether is an error born chiefly of ignorance. But what about Japan? Sir Douglas Hague's original protest raised this question but without conviction. It is, of course, a red herring, as the Japanese now seeking to legislate for more creativity within their system would be the first to confirm. Japan has done well, over the years, by also adapting technology from elsewhere to modern market needs, often the needs of markets elsewhere. But even Japan is not sure of sustaining its technological growth beyond the point reached by others with its own unembellished basic research. And while the Hagues of this world may understand the subtlety of the problem, they should choose their words more carefully so as not to give a licence to those who do not or prefer not to understand. In Britain just now, there are too many philistines of that kind roaming the streets and even the corridors of Whitehall. □



Dollar fragility

*Why must researchers worry about the dollar?
Because research funds may vanish.*

LAST week's predicted further fall in the value of the US dollar is not the first this year, and will not be the last. Last month's tumble was attributed to the financial difficulties of several Ohio semi-banks and last week's to the estimate of slower US growth this year (1.8 per cent) than last (6.5 per cent), but these are occasions, not causes. The fragility of the dollar is a consequence of the US government's huge budget deficit and of the certainty that even if the administration and the Congress agree that it should be cut, the effects cannot appear until the beginning of the next financial year in October. Can the world wait that long?

The fear that time is running out is barely concealed. Even finance ministers at last week's meeting of the International Monetary Fund in Washington allowed their determination not to be seen to rock the boat to be compromised by anxiety. What worries them is that the US administration is now a tight box. Either it will have to cover its deficit by increasing interest rates still further or give up trying, print money instead and put up with the inflation that would follow. The explanation that last week's fall was triggered by people's calculations that the administration would now try to stimulate the economy by reducing interest rates is not to be taken seriously, for cost-free options have gone the way of the free lunch. Nor does it make sense to ask, as some do, that other industrialized nations should assume the burden of stimulating economic growth by government policy, for inflation-free growth outside the United States will be possible only when the flow of funds to the United States in the past few years has been substantially redirected, adding to the problems of the dollar.

This is why talk of monetary reform has become fashionable again. Even Mr James Baker, the new Secretary of the US Treasury, is in favour. The trouble is that nobody knows what should be done. Things have changed a lot since John Maynard Keynes went off to Bretton Woods in 1943 and drafted the monetary rules by which the world lived for the succeeding thirty years. The essence of Keynes's recipe was a blend of currency stability by means of declared exchange rates and a mechanism for the supply of credit, through the International Monetary Fund. One change since then is that there is now almost a surplus of international credit, much of it stimulated by the steep increase of the price of oil in the 1970s, whose movement is uncontrolled and hard to anticipate. Another is that the financial markets have become technically more efficient and sophisticated than Keynes could have imagined. The needs now are to find ways of giving currencies a measure of stability even when they are traded freely, and to shield developing countries from the damage they can suffer from these processes. After a decade in which most governments have been indifferent to these needs, it is gratifying that most governments now accept the need for reform.

Part of the trouble with the present system is that huge sums of money can slop about within it, in pursuit of the marginally different interest rates on offer at different places. The volatility of the system would be much reduced if only governments would agree to limit the proportion of their borrowing that consists of short-term debt. (In a typical week, the US Treasury now raises between \$10,000 million and \$20,000 million of short-term money, most of it to cover maturing debts and not the underlying deficit.) Self-denial like this by governments would not be very different from what governments require of their commercial banks. In the process of reform, Europe has a particular responsibility (and opportunity). A more thorough coordination of European arrangements for coordinating monetary policy would increase the inertia of the system without, in the long run, biasing the ways in which money moves. The European Monetary System, which ties together major European currencies (but not sterling), is a start but an imperfect one, while it will be decades before the European accounting unit (called the ECU) will become a currency. One objective, now, should be to set up a pan-

European central bank with the funds and the independence to be effective.

The monetary problems of the developing countries are more intractable, partly because most governments do not accept the logic of the need to act. Why else would so many rich countries still restrict the import of simple manufactures, but not the raw materials on which many developing countries are too dependent for what passes for prosperity? While restraints like this persist, nobody should be surprised that developing countries cannot easily attract the funds they need for development (although they are also, often, their own worst enemies by the political hazards with which they confront potential investors). In this sense, last week's meeting of the International Monetary Fund at Washington must know that its talk of insuring private companies' investments in developing countries, by some mechanism yet to be devised, is a palliative, not a solution. The monetary problems of the poor will be solved permanently only when the industrialized world accepts that they have a right to earn their way out of trouble, and is then prepared to help them do so through agencies such as the World Bank, designed for just this purpose. □

Tarnished initiative

The offer to allies of a share of the US defence budget has been indelicately handled.

THE second subject (monetary policy is the first) that will preoccupy the summit meeting of the major powers due next month is that of what should be done about the US plan to build a defence against ballistic missiles, sometimes known as the Strategic Defense Initiative, but which is otherwise star wars. European governments represented at the meeting will be in a fix. Sceptical of the scheme from the outset on the grounds that it would probably be ineffectual and that the development of such a system would cause more international trouble than even success would be worth, and in some cases fearful that success would lead to a US abandonment of Europe, they have been tempted against their better judgement into collaborating with the planned technical developments (see p.660). For many, the principle seems to be that they will then at least know what is going on.

The US government's position is not very different. There was nothing in President Reagan's speech in March 1983 to suggest that star wars was intended to be collaborative venture, or even that the governments of Western Europe could fail to share the president's own sense of wonder at the supposed benefits of a world (or part of it) blessed with immunity from nuclear attack. In the event, the scepticism of European governments blended with the administration's own sense that there may be safety in numbers in the trying months ahead have combined to persuade the United States that collaboration would be worthwhile. The reluctance, or surprise with which the administration has come to this view is clear from the letter from Mr Caspar Weinberger to the political heads of potentially collaborating countries last month: in these days of word processor, it is surely needlessly offensive to send what seems to be a circular letter, in which recipients are invited to render "your nation's" response.

As things stand, there is no way telling what European states should say. The offer of collaboration is too innocent of detail to be appraised. The governments that fear that collaboration will be a formal matter only may be right. Those who hope to make money out of the scheme may, alternatively, hit the jackpot. But nobody can know for sure — and certainly not within the 60 days first specified in Mr Weinberger's letter. But the simple rule should be this: if collaboration is genuinely intended, it should be complete, and should allow collaborating European partners to influence not merely the development of star wars technology but also the use that is made of it. And otherwise, if collaboration is offered on terms that are less than full, it should be politely declined on the grounds that it cannot fail to be comprising. That is one version, but that there are others which are less polite is what some governments will be saying in Bonn. □

AIDS progress

Synthetic vaccine only a distant prospect

Atlanta, Georgia

THE virus responsible for acquired immune deficiency syndrome (AIDS) in humans may have entered the population recently by transmission from monkeys, according to a study of African green monkeys by Max Essex of Harvard University. Essex has found that apparently healthy monkeys show strong antibody recognition of STLV-III, a virus associated with immune deficiency disease in rhesus macaques which has many similarities with human AIDS virus. A simian origin for the virus would be consistent with new data suggesting that AIDS may have been present in humans of central Africa since before 1973.

Essex reported his findings here at the first international conference on AIDS, attended by two thousand researchers from 26 countries. The meeting was sponsored by the US Department of Health and Human Services and the World Health Organization, among others. The availability of two different primate models of AIDS, one with disease and one without, should be invaluable for development of an AIDS vaccine: comparisons between viruses might indicate those conserved regions that might be most usefully imitated in a synthetic vaccine produced by genetic engineering, so avoiding the problems caused by antigenic variability.

A vaccine that would prime the immune system to deal with the AIDS virus is one of the best hopes for dealing with the disease; it seems unlikely that a treatment will be found that will eradicate the infection once it becomes established in an individual, where DNA copies of the virus are incorporated into the host's own DNA. James Curran, head of the AIDS task force at the Center for Disease Control (CDC), said that if a vaccine became available it might be necessary to vaccinate all US citizens, starting first with those in high risk groups who at present show no evidence of infection. A vaccine is still several years away, however.

The latest data on the spread of AIDS in the United States are grim but unsurprising. By last week 9,608 cases had been diagnosed of which 4,712 have been fatal. Estimates of the number of US citizens who have already been infected with the AIDS virus, as indicated by the presence of serum antibodies, range from 0.5 to 1.0 million; more than 10 per cent of these are likely to develop the disease over the next five years, even if all further transmission of the virus stops immediately. The doubling period for the number of reported cases in the United States has increased slightly, but is still around 8-9 months. Principal con-

clusions from the conference have been reported to the US Department of Health and Human Services, and a separate report is being prepared for the World Health Organization.

New estimates have been made of the latency period between infection and the onset of disease symptoms, and it seems that the mean latency may be five years or even longer; this suggests that many people in high risk groups may already have been infected but are clinically unaware of the fact.

Between 70 and 90 per cent of haemophiliacs who receive regular infusions of factor VIII concentrate have detectable serum antibodies to the AIDS virus, and some data suggest that more than five per cent of these are likely to develop AIDS, with a further number developing "lesser AIDS" — AIDS-like symptoms that nevertheless fall outside CDC's definition of AIDS (Elaine Eyster, Pennsylvania State University School of Medicine). Estimates of the proportion of people with antibodies who will develop the disease are likely to go on increasing. Among one group of New York promiscuous homosexuals, the proportion of those with antibody that have developed AIDS is as high as 20 per cent (J Goedert, National Cancer Institute); it is not yet clear whether this high figure is due to some complicating lifestyle factor or whether it is because the virus has been present for longer in that group.

Very preliminary data from wives of haemophiliacs suggest that the AIDS virus is transmitted heterosexually by repeated contacts from male to female (J. Goedert, National Cancer Institute). A claim for female to male heterosexual transmission from prostitutes was also made (R. Redfield, Walter Reed Army Medical Center), though these data are also equivocal.

Some modestly encouraging results for treatment of AIDS with anti-viral drugs were reported, although none amounted to a major breakthrough. It is now apparent that neurological symptoms such as dementia are present in high proportion of AIDS patients, often associated with encephalitis.

US Secretary of Health and Human Services Margaret Heckler told the conference that AIDS is the chief health priority of the administration, but her speech, which studiously avoided mentioning the principal risk group, male homosexuals, did nothing to soothe the anger of homosexual activists and AIDS patients attending the conference.

Robert Gallo of the National Institutes of Health, who first isolated the virus known as HTLV-III, made a plea for less petty nationalism in the competition be-

tween research groups in different countries. Gallo's plea was not immediately effective: the AIDS virus continues to be referred to by at least three different names depending on the affiliation of the speaker, despite agreement that the virus isolated by different groups is basically the same, and despite an informal agreement to call the virus HTLV-III/LAV (or LAV/HTLV-III) in recognition of the work of the French group under Luc Montagnier at Institut Pasteur in Paris. The name issue is to be tackled later this year by the International Committee on Virus Nomenclature. Gallo also took many by surprise by calling for an end to the cult of the instant media authority.

All the blood banks in the United States are now screening donated blood for AIDS virus using enzyme-linked immunosorbent assay test kits manufactured by commercial companies, but implementing the tests has not been without its problems. According to Gerald Sandler of the American Red Cross, an anomalously high proportion of samples tested with the Abbott Laboratories' kit produced evidence of AIDS antibody on first testing in some areas of the country, although the proportion of repeatedly reactive samples was close to the expected 0.2 per cent. Samples that produce positive results on two subsequent tests are confirmed by Western blotting, but it is apparently impractical to carry out Western blots on all samples that produce a positive result just once.

It is also becoming clear that after infection by the AIDS virus, several weeks or months may pass before antibodies reach easily detected levels; any test that looks just for antibodies can therefore be at best an imperfect protection for blood banks. In Britain, initial investigations of the Abbott test have produced disappointing results, and it seems most likely that Britain will develop its own test for blood banks rather than importing US products; price is also a factor.

There is still no agreement on whether US donors who show clear presence of antibody that is confirmed by Western blot should be told of the results of the test. The American Red Cross says it is still reviewing options, and organizations in different states are following different paths. California has adopted a moratorium on conveying test results to donors until alternative testing centres are established where people can be tested for antibody without giving blood.

The fear is that as long as testing is available only at blood collection centres, people in high risk groups will be attracted to offer their blood in order to be tested: questionnaires indicate that most people in high risk groups do want to know whether they have circulating antibody, despite the uncertain prognosis. Some federal funds are likely to be made available to establish several hundred alternative test centres, though the amount is still being argued.

Tim Beardsley

Star wars

Israel attracted by spin-off

Rehovot

ISRAEL is almost certain to accept the invitation it has received from the United States to participate in the Strategic Defense Initiative ("star wars") research programme.

If the government does so, it will enjoy the support of most of the scientific community, which hopes that the project will provide research funds that are no longer available from other sources and also create positions for young scientists who might otherwise leave the country.

Professor Josef Singer, president of the Haifa Technion and an aeronautical engineer who has himself done defence research, supports Israeli participation but warns his colleagues that they should not expect the United States "to spend billions of dollars and to engage thousands of scientists and engineers".

Singer points out that the US armed forces have already spent millions of dollars over the past 30 years in support of research projects at institutes in Israel. "And", he says, "while Israel stands to gain more in resources and know-how, it also has a contribution to make. Otherwise, it would not have been invited to participate." He expects there to be benefits for Israel in the high-technology industries. And off the record, other Israeli scientists say that star wars research, although primarily aimed at finding a way to bring down Soviet ballistic missiles, may help to improve Israel's military potential closer to Earth. Specifically, they speculate about the development of an innovative deterrent that might compensate for the fact that shortage of money and manpower make it difficult for Israel to match the military build-up of its Arab adversaries.

A dissenting opinion comes from Professor Giora Shaviv, an expert on space research at the Haifa Technion, who dismisses the invitation to Israel as a propaganda move and says that Israeli researchers "have no real role to play" in the star wars scheme. A small number of politicians and academics have raised the subject of the moral and political implications of Israel's participation in the US programme.

Meir Stiglitz, who teaches a course on nuclear weapons at the Hebrew University's International Relations Department, claims that Israeli leaders are so blinded by the immediate benefits of cooperation with the United States in this sphere that they overlook the potential long-range dangers. As he sees it, these include the possibility that Israel may become a "desirable and legitimate" target for Soviet nuclear missiles, and might be attacked in order "to send a message" to the United States.

Nechemia Meyers

Star wars

France rallies European forces

THE French government is determined to shake Europe awake, and to encourage all the nations of Europe to take a great leap forward in science and technology, by creating a new European research organization with considerable powers and funds, dubbed "Eureka". Or it may be that the French government is scared stiff of participation in the US Strategic Defense Initiative (SDI, or star wars) because it feels this might suck France dry of some of its best weapons technologies, and wishes to form a common European industrial and research front. Eureka would be that front.

Which is the truer picture of French intentions? Probably both are true, which has made it a little difficult for Western European foreign ministers to respond to an urgent letter, proposing the flexible and rapid creation of Eureka, addressed to them last week by their French opposite number, M. Roland Dumas.

In the letter, Dumas says that Eureka (which in a rare French linguistic concession he gives the English name "European Research Coordination Agency") would set up major programmes half-funded by Eureka and half by industry in optronics, new materials, high-powered lasers, artificial intelligence, high-speed and ultra-miniaturized microelectronics and space. Any European state should be able to join in any programme, Dumas suggests, leading to a "variable geometry" Europe for research. But what research? Dumas' shopping list is remarkably similar to one that might be drawn up by SDI, rather than the new technologies in general (where is biotechnology, for example?), but the letter does not mention SDI by name.

Nevertheless, French interest in creating what the Prime Minister, M. Laurent Fabius, calls "a European area for research" is well known, and long pre-dates interest in SDI. France has doubled its government civil research spending since President Mitterrand came to power in 1981, but now realizes that France cannot go it alone against the United States and Japan. Thus Pierre Papon, director-general of the principal French research council, the Centre National de la Recherche Scientifique (CNRS), which has 10,000 researchers, made a detailed tour of Europe last year seeking bilateral research agreements, and the present research minister, ex-president of the French space agency the Centre National d'Etudes Spatiales (CNES), has also shown himself a dedicated European.

But these individuals and others in the French science political scene may be getting impatient at the pace of European integration in science and technology. ESPRIT, the European programme of research in information technology, flagship of the European Economic Community (EEC) programme in

integrated "pre-competitive" research, has been extremely slow to spend real cash, and for the moment can claim only that it has brought a few European companies to the same table. The only really working institutions of any size seem to be the European Organization for Nuclear Research, CERN, which has only a moderate economic impact, the European Space Agency (ESA) of which Curien has been chairman, JET (the European fusion experiment, with impact only far in the future) and the European Airbus.

France wants much more, and earlier this month Jacques Delors, the new French president of the European Commission (the EEC bureaucracy), tried to provide it. In a written presentation, he asked European ministers at a summit meeting to double EEC research spending from its present three per cent of the Brussels budget to six per cent, and he linked the increase with the need for Europe to respond to SDI. Ministers — who at present in the 10-nation EEC must agree any proposal unanimously — threw out both ideas. Many states were, it seems, prepared to increase the research budget, but by nothing like the factor of two, and the SDI link was described by some participants as "half-baked", rushed and ill-prepared. There seems to have been no collaboration between Delors and the French government over the issue, and it may be that Delors had had wind of the Eureka proposal and wished to put up an EEC stake, fearing the likely weakening effect of Eureka on Commission power in research.

Nevertheless, Delors was unsuccessful, and now Dumas has proposed Eureka, determined, it seems, to see something happen for European research. But the announcement of his initiative last week in Paris was not made through the usual channel, the Wednesday council of ministers meeting, but direct from the ministry of foreign affairs itself. Staff of CNRS "read about Eureka in the press". There seems to have been "a little problem of communication" between the foreign ministry and the research council, though Pierre Papon "has since become involved in discussions".

Moreover, the links with defence have clearly become stronger, as Dumas' letter to foreign ministers was delivered a few days in advance of this week's meeting of the Western European Union, a recently-revived forum in which seven of the major European states discuss defence matters. The letter is intended for discussion "on the margins" of the meeting. Dumas will no doubt be expressing concern about French industry, in which spending on armaments has become the principal component. Last year, for the first time, France exported more arms to the United States than it imported from that country. And

recently a number of French companies including Thomson, the nationalized electronics company, have gone well into the black through trade in armaments, much of it to the Middle East. Often this trade relies on high technology weapons, such as Exocet, the missile which was so dangerous to British ships in the Falklands conflict. France does not seek a one-way drain of expertise or experts to the United States through premature involvement in the star wars programme.

Thus Dumas is concerned by the position of Britain, where foreign minister Sir Geoffrey Howe has criticized SDI so strongly (on strategic grounds) that France fears Britain may not participate at all. Dumas is slightly happier with West Ger-

many, with which France has increasingly been developing a special relationship in Europe. There, Chancellor Kohl has expressed full support for Eureka but has gone further than the French by agreeing separately with the United States to take part in the SDI research programme. Kohl will send a team of negotiators to Washington to discuss the matter, he has told the Bundestag.

Some united participation in SDI is what France wishes to see. And it is even preparing its public SDI is "star peace", not "star wars", defence minister Charles Hernu claimed on French radio last week. "And France must have its place in this star peace", he proclaimed.

Robert Walgate

Star wars

US pressure on Japan

Tokyo

THE US Strategic Defense Initiative ("star wars") seems to have found a fan in Japan's Prime Minister Yasuhiro Nakasone, who has been extolling its virtues in recent Diet speeches after his meeting with President Ronald Reagan in Los Angeles earlier this year. Pressure for Japan to make its support official is being kept up by a series of visits from US experts. The real issue, however, is how far Japan is prepared to move from its pacifist constitution, which prevents participation in collective military efforts.

At the end of last month, US Defense Secretary Caspar Weinberger formally asked Japan as well as other US allies to inform the United States within 60 days whether it is willing to participate in the US research programme for the Strategic Defense Initiative (SDI). Weinberger's request was immediately followed up by a series of visits to Japan by US military experts and a delegation from the North Atlantic Treaty Organization (NATO) to "explain" SDI and to look for areas of cooperation.

Although the 60-day deadline has subsequently been dropped by the United States as a bureaucratic "error", further teams of US experts are expected in Japan, starting this week.

Japan's space technology is of course still relatively backward and largely developed under US licence. But where Japan has something to offer is in its electronic communications technology, particularly in the use of extremely-high-frequency microwaves to transmit huge quantities of data from satellites to the ground.

Japan's defence agency has, in fact, already completed a study of SDI aimed at assessing the efficiency of both chemical and nuclear-pumped X-ray lasers. Although the study seems to say only that enormous improvements in aiming accuracy would be required to shoot down intercontinental ballistic missiles, that it was carried out at all is an indication of the

extent of Japanese interest.

In theory, Japan should have no military forces at all; Article 9 of the constitution (forced on Japan by the United States at the end of the Second World War) says that "land, sea and air forces, as well as other war potential, will never be maintained". In fact considerable "self defence" forces have been in place for years. But until very recently the Japanese government has regarded supplying military technology to foreign countries as going much too far. In 1983, however, considerable pressure by the United States forced Japan to accept a bilateral agreement allowing the transfer of military technology to the United States — again, US interest lay largely in the communication field.

Star wars would be another step up in Japan's military involvement with the United States, but would be seen as a step onto a slippery slope by opposition socialist parties. Government representatives have already had to resort to semantic juggling to make a possible involvement respectable. Military technology transfer is as yet possible only to the United States, but SDI would probably involve other countries. The Foreign Ministry's Treaties Bureau chief Hisashi Owada sidestepped this issue by saying Japan can help the United States with the research, "even if NATO countries take part in it, so long as military technology supplied to the United States remains there".

If SDI research does go ahead, it will bring closer the day when Japan's giant corporations can go into the military export business. But Nakasone may find there is more opposition to the plan from the general public than he has bargained for. Like other world leaders, he has found it easier to agree with Reagan when having a cosy face-to-face chat than to sell Reagan's ideas back home. Nakasone is scheduled to have his next get-together with Reagan on 2 May, just before the Bonn summit, and star wars is high on the agenda.

David Swinbanks

Spain

Law of Science brings change

Barcelona

SPANISH research institutions are likely in the next few months to face a similar upheaval to that experienced in the universities as a result of the Law for University Reform (see *Nature* 312, 8; 1984).

The Spanish government has now approved the Law of Science (Ley de Fomento y Coordinación General de la Investigación Científica). The bill has been sent to parliament, where the Socialist Party has an absolute majority in both houses, and a vote is expected before the end of the year. Government approval of this law, after months of discussion between different ministries, may represent an important step in the reform of the organization of research in Spain.

The Law of Science aims to establish mechanisms to define research priorities, to programme the resources for research and to coordinate the work of those in universities and research institutes. The bill also aims to redefine the role of the main public bodies concerned with research, which until now have often had overlapping functions because of the accumulation of 50 years of contradictory legislation.

Under the new law, the National Plans for Research will define research priorities and determine how public funds should be distributed. The draft of a National Plan for Research and Development for the next five years has also been presented as an annex to the law. Subjects to which it would give priority include microelectronics, biotechnology, transportation technology and high-energy physics. Fellowships tenable both in Spain and abroad are intended to double the number of researchers in industry and public organizations within five years.

The Law of Science also proposes a new structure for the main public research bodies, including CSIC (Consejo Superior de Investigaciones Científicas) and JEN (Junta de Energía Nuclear), to be called the Centre for Energy, Environment and Technological Research. They will be organized so as to promote the mobility of staff between different institutions.

The need for a change in the regulation of Spanish research has been generally acknowledged. Officials in the Ministry of Education and Science believe the new law may help to break down the barriers between people doing the same work in different departments, to promote interaction between public and private research and to continue the trend towards better scientific control and less bureaucracy in the administration of public resources. If these reforms are coupled with increased public and private spending on research, the law will enhance Spain's contribution to science and technology.

Pedro Puigdomenech

Cancer research

France moves ahead

FRANCE is moving ahead of Britain in cancer science, according to a literature review* carried out by Jean-Claude Salomon and a colleague at the Institut de Recherche Scientifique sur le Cancer at Villejuif near Paris.

Salomon based his studies on the 300 or so key papers selected annually for summary in the *Year Book of Cancer*, a regular series of reviews of the world cancer literature produced for the education of US doctors. The yearbooks, edited at the M.D. Anderson Hospital and Tumor Institute in Houston, Texas, "satisfy certain criteria of rigour and objectivity", says Salomon, "though they may perhaps favour American authors". The editorial board of 155 discipline specialists contains representatives from 14 countries, but most members of the board are American. Each year, this board selects the most significant papers from among around 50,000 published each year in 1,200 major journals for summary in the yearbook.

Using this source as a measure of international productivity of cancer science (including clinical papers), Salomon discovered that from 1965 to 1982, the proportion of papers summarized in the yearbooks that were British had declined from 8.3 per cent (the then highest figure outside the United States) to 4.2 per cent, whereas French articles had climbed from 2.3 per cent to 4.6 per cent, putting France in the role of leading nation for cancer research outside the United States.

The founder-editor of the yearbooks (and founder of the M.D. Anderson Tumor Institute), Dr Randolph Lee Clark, believes these figures may well represent real trends. While France had seen fit to invest in 25 "cancer centres" during the period reviewed, the British scene had been almost stagnant, he observed.

Salomon's figures also show a poor result by West Germany and by Israel, each with under one per cent representation in the yearbooks.

In France itself, the clearest signal from Salomon's review is that there has been a strong shift in cancer science from Paris to the regions. Between 1965 and 1976, three-quarters of French articles summarized in the yearbooks came from Parisian institutes. But from 1978 to 1982, Paris could claim little more than half, a result with statistical significance at the one-in-a-thousand level, according to Salomon. In particular, the renowned complex of institutes at Villejuif, including Solomon's own workplace, which had more than 40 per cent of the articles in the first period, had only 12 per cent in the second. This may be a measure of the effectiveness of a big French effort at decentralization 12-15 years ago.

Could there be bias in the selections? Dr Clark claims not, and that he tries to en-

sure that "there are no national barriers" in the coverage of the yearbook. The sole British representative on the editorial board, Dr Neils Atkin, cytogeneticist at

Heavy ions

Darmstadt to leapfrog Berkeley

Hannover

THE heavy-ion laboratory at Darmstadt, south of Frankfurt, last week received its reward for its discovery of element 109 three years ago — DM275 million (£65 million) with which to build a bigger and better ion accelerator.

The West German federal minister for research and technology, Dr Heinz Riesenhuber, announced on Thursday that the laboratory would be allowed to go ahead with a plan it had nurtured for at least the past four years. The scheme is to feed ions from the existing linear accelerator UNILAC, at an energy of up to 20 MeV per nucleon, into a synchrotron designed to yield energies 100 times as great. There is also to be a storage ring equipped to bunch and to homogenize the velocity of heavy ions from the synchrotron.

Those working at the laboratory are ecstatic about the prospect. While the maximum energy from the new machine is not very different from that of the similarly constructed BEVALAC at Berkeley in the United States, the intensity at Darmstadt will be greater by a factor of 1,000.

The search for other exotic nuclei continues at Darmstadt, even though people talk less confidently now than a few years ago of a possible "island of stability" in the region of atomic number 116 where

Mount Vernon Hospital, Middlesex, described Salomon's observations as "rather sad" for Britain and added that "I can't think of an explanation that would give us any relief". The editors try to be objective, he said. **Robert Walgate**

* Salomon, J.-C. & Blanchin, P. *Bull. Cancer (Paris)* 71, 382-387 (1984).

heavy nuclei would be relatively stable. Indeed, the existing linear accelerator has just finished a two-week run looking for signs of element 110 in the collision of nickel and lead ions. But people now insist that the intended machine will be justified even if no further heavy elements are found.

One field of interest is the beta-decay of heavy nuclei stripped entirely of their orbital electrons. Collisions at 2,000 MeV per nucleon may also simulate the condition of nuclear matter, as in neutron stars, while there is virtually the whole of the atomic structure of heavy ions with many charges to be explored. On the technical side, there are interests in the use of heavy-ion beams for pumping X-ray lasers, and for investigating the properties of dense cold plasmas (which may be crystalline).

The plan is that the extension of the Darmstadt laboratory should be completed in the next five years. The laboratory hopes to increase its staff (now 450 with 50 supernumeraries), although not by the whole of the 70 positions asked for. Of the extra funds now made available, 90 per cent will come directly from the Federal Ministry of Research and Technology, while the remainder will be divided (in the ratio 9:1) between the federal government and the Land of Hesse. The laboratory's budget is now DM 80 million a year.

John Maddox

Director for Newport News facility

Washington

HERMAN Grunder of Lawrence Berkeley Laboratory in California has agreed to assume on 1 May the directorship of the Continuous Electron Beam Accelerator Facility, the machine that may be used to study "bags of quarks" and that is to be built in Newport News, Virginia, by a consortium of south-eastern universities led by the University of Virginia.

Grunder's decision to accept the position is seen as a major step toward legitimizing the project, which has been plagued with political and technical problems from its start two years ago.

First there was a major political battle with Argonne National Laboratory over who should build it. Although the Department of Energy Nuclear Science Advisory Committee (NSAC), traditionally the last word of the physics community, endorsed the Virginia group's proposal, Argonne

broke protocol and appealed directly to the Secretary of Energy, enlisting the aid of the Illinois congressional delegation. In the process, enough questions were raised about the technical expertise of the Virginia group, the technical feasibility of its design, and the accuracy of its cost estimates (which have already grown from the \$100 million figure presented to NSAC to more than \$200 million in congressional testimony last year) that a sceptical Congress insisted on a go-slow policy. Construction funds requested in the very first year have yet to be approved. The resulting lack of confidence in the physics community that the project would ever in fact begin made recruitment of a director difficult.

Meanwhile, though, the accelerator received a much-needed boost when NSAC, asked by Congress to re-evaluate the project, gave it a firm reendorsement.

Stephen Budiansky

Bhopal aftermath

Legal complications mount

Washington

THE lawsuit filed in New York two weeks ago by the government of India on behalf of victims of last December's disaster at the Bhopal Union Carbide plant has placed on the table a complex series of legal issues that will have to be resolved before even an out-of-court settlement can be reached.

Two of the thorniest questions are whether the US district court will allow the case to proceed in the United States and, if so, whether Union Carbide will be able to countersue the Indian government. To be resolved, too, is the validity of India's claim to be the exclusive representative of the victims and the status of nearly 1,000 lawsuits filed by victims in Indian courts.

In preliminary hearings last week, Judge John Keenan refrained from showing his hand on any of these issues. But, invoking the broad discretionary powers accorded federal judges in the United States, he urged Union Carbide to make an immediate contribution of \$5-10 million for emergency relief; Union Carbide the next day agreed to turn over \$5 million. Both Keenan and the company made clear that this was not an admission of liability (or even of the US court's jurisdiction).

From the start, the Indian government has said it intended to sue Union Carbide in the United States. Although any compensation awarded by US courts for lost wages, for example, would be based on Indian standards, US courts are much more generous in their compensation for less tangible losses, such as pain and suffering, than are their Indian counterparts.

Awards in US courts are set by juries, usually more sympathetic to victims; Indian cases are heard by judges only. And Indian judges have been very reluctant to award punitive damages, except in cases of clear deliberate wrongs. US courts generally allow punitive damages against defendants who have shown reckless disregard for safety, even when not deliberate.

If Keenan agrees to hear the case, the pressure on Union Carbide to settle out of court will increase measurably. The company has already made one offer of settlement, unofficially put at \$200 million, an amount reportedly equal to its insurance coverage. The Indian government has rejected that offer.

If the case is tried in the United States, the company would be facing the real possibility of a judgement of perhaps \$1,000 million, approximately one-tenth Union Carbide's entire net worth. The pressures on the company for a settlement are already considerable in any case.

But meanwhile Union Carbide is applying some legal pressures of its own to urge concessions on the other side. The company's lawyer has made it clear that he will argue that the case should be heard in India as the more appropriate forum. He

has also suggested he may move to have the US law firm retained by the Indian government disqualified because it once represented Union Carbide, though on an unrelated matter. And there is also the threat of a countersuit against the Indian government. Normally, governments enjoy "sovereign immunity", and cannot be sued for their actions. But under US law, a foreign country that brings suit in a US court cannot claim sovereign immunity in defence against counterclaims. Union Carbide would probably argue that by allowing squatters to settle near the Bhopal

plant in defiance of municipal plans, the government shares responsibility for the injuries.

A wild card in any resolution of the claims is who has the right to represent the victims. The Indian government has enacted a law declaring the government the exclusive representative, while allowing victims to hire their own lawyers who could join the government's action. One faction of the US lawyers who signed up victims — a faction that has been hoping to organize the case into a class action — is challenging the Indian law in Indian courts. It is being joined there by Indian lawyers whose suits filed against Union Carbide in India have been frozen by the law.

Stephen Budiansky

Handicapped infants

Physicians win the day

Washington

THE furore that greeted the Reagan administration's regulations for the treatment of handicapped infants seems finally to have been quelled. Last week, the Department of Health and Human Services (HHS) issued new rules that for the most part defer to the physician's judgement in deciding how to treat such infants and that keep the federal government out of the enforcement business.

The new rules, the result of an arduously crafted congressional compromise last fall, require only that each state's child abuse agency should develop procedures for investigating and "responding" to reports of medical neglect of infants and that they have the legal authority to seek a court order to prevent "withholding of medically indicated treatment". The rules do not apply to cases in which the treating physician believes, in his "reasonable medical judgment", that:

- the infant is in an irreversible coma;
- treatment would merely prolong dying, fail to correct all of the infant's life-threatening conditions, or otherwise be futile; or
- treatment would be virtually useless and would itself under the circumstances be inhumane.

Earlier versions of the rules stirred great resentment among paediatricians and hospitals. Under the original rules, issued in the spring of 1983, hospitals were required to post prominent signs warning that "discriminatory failure to feed and care for handicapped infants in this facility is prohibited by federal law".

The rules also established a "Baby Doe hotline" for receiving reports of medical neglect; HHS set up an investigative team that would be sent out on short notice to examine these reports. The hotline received about 400 calls, most of them from cranks; in the four cases examined no violations were found. But critics complained that the parents of the infants and the hospitals had been put through a needless ordeal. The regulations had been in effect for only a

month when they were thrown out by a federal judge acting on a lawsuit brought by the American Academy of Pediatrics and others.

The physicians' organizations also argued that the two well-publicized cases were carefully scrutinized by state authorities, and in both cases local courts refused to intervene. Almost all states include medical neglect in their definition of child abuse; courts have often intervened in cases where parents have refused to allow treatment for their children on religious grounds. The physicians' group thus argues that no federal involvement is necessary.

The new rules indeed go a long way to acknowledging that, and make clear that the federal government is not to establish standards prescribing appropriate treatments. In ordering HHS to issue these rules, Congress in fact seemed chiefly concerned with defusing the issue in such a way that the administration — and the right-to-life groups that the administration had been appealing to in issuing the original rules — could save face. Physicians' organizations remain concerned, however, that even the new rules will put improper pressure upon paediatricians always to choose heroic, aggressive treatment for fear of being accused of withholding treatment.

HHS says it received 116,000 letters commenting upon a draft version of the new rules, published in December. One major change that HHS made from the draft to last week's final version was to remove a set of definitions of terms from the body of the rules into a non-regulatory appendix. Physicians said the definitions of such terms as "life threatening condition" included specific examples and placed the federal government in the position of dictating specific procedures. The six principal Senate sponsors of the compromise bill urged the agency to consider these complaints seriously, noting that there was a danger of undoing the careful balancing that went into the bill.

Stephen Budiansky

US universities

All change at St Louis

St Louis, Missouri

THE Washington University, part of this city's pride (after the arch across the Mississippi and its baseball team, the Cardinals), is part way through an administrative revolution that keeps many of the faculty awake at nights. Two years ago, the trustees decreed that the School of Arts and Sciences should follow the professional schools into financial and administrative autonomy. One result could be that academic departments less able to fend for themselves financially — philosophy but even physics — might go out of business.

The School of Arts and Sciences has had autonomy thrust upon it. Two years ago, it was the only school of the university not paddling its own canoe. The argument for change starts from the observation that other schools that have been independent for some time (medicine conspicuously) have prospered, partly from the ease of making deals with outside backers and partly by commanding the loyalty of their alumni. But then, as the administrators are fond of saying, is this not what the Harvards of this world do?

Under the arrangement, called "being put on reserve" in reference to the share of the whole endowment that autonomy commands, a school is the fiefdom of its dean. Policy on admissions, curriculum and even salary levels is determined separately. Now, in the School of Arts and Sciences, tuition fees (in excess of \$8,000 a year for undergraduates) go into a pool with other income, but the school pays a levy to support central activities.

Worries among the faculty take several forms. The philosophy and classics departments (large and small respectively) may not be financially viable, but can a university claim the name without offering courses in such fields? But even physicists fear that the engineering school, which now pays a fee to arts and sciences for service physics teaching, may decide to teach its own physics as an economy. (The chancellor, Dr John Danworth, says the service fee, fixed by the trustees, could not be changed without their consent.) Then there is the general worry that autonomy will turn all academics into fund-raisers, or that only successful fund-raiders will prosper.

The chancellor is anxious that autonomy should work, but is for the time being open-minded about the change. If it does not bring the benefits intended, the trustees would think again. And a successor might wish to see the omelette unscrambled. But in Danworth's view, only autonomy can make a faculty responsible for its academic future. He cites his own earlier success, as head of the medical division, in turning a bankrupt dental school into a successful enterprise with the help of a tapering subvention from the centre (\$600,000 over three years) coupled with the threat that

failure would spell closure.

In any case, Danworth says, it is no longer possible for the head of a private university, forever occupied with fundraising and negotiations with government agencies, to be so involved with academic life as to provide single-handed academic leadership. This leads him, after careful thought, to an unyielding position on the issue of the weaker departments; if the faculty of arts and sciences should think of closing one of them, the whole value of independence would be lost if he or the trustees should intervene to halt the closure. But what a chancellor can do is to fire a dean who has failed to carry the faculty with his decisions. Danworth is said to read faculty meeting minutes like a hawk.

Ultimately, the success of the experiment will hang on whether arts and science can raise the funds needed to stay alive, and on whether the faculty is sufficiently committed to the ideal of a liberal education for the well-to-do to indulge their less commercially viable colleagues.

That money is so dominating surprises nobody at St Louis, although most academics are more concerned with the competition for high-quality students, which they say is a prerequisite for commercial success. They point to rising scores in the national college boards examinations and a widening catchment area (with only 20 per cent of undergraduates from the locality). Physics has now followed engineering in offering a handful of undergraduate scholarships, saying that even students who compete unsuccessfully for them leave their applications in. The faculty also claims that other universities are prevented by the egalitarian climate

from following this elitist course.

So far, the university has been lucky in its donors. The chemical company Monsanto, only one of the reasons why East St Louis (across the river in Illinois) is such a slum, made a 5-year grant of \$27 million to the medical school two years ago, from which non-medical biologists also profit. McDonnell Douglas, the aircraft manufacturer, has been generous over several years, partly because of the interest of the late founder of the company (known generally as "Mr Mac") in space research (whence the flexibly endowed McDonnell Center for Space Research) and in parapsychology (engendering less pride). By contrast, General Dynamics, also based in St Louis and now in trouble for misdemeanours such as lodging a senior manager's dog in a commercial kennel at the Pentagon's expense, is almost conspicuous for the modesty of its support.

As a research university, Washington dates its rise to the post-war chancellorship of A.H. Compton, who arrived at St Louis in 1945 with a group from Los Alamos. More recently, the medical school has made the running.

For a big-city urban university, Washington is surprisingly well-provided. It is compact (mostly on one campus, with the medical school two miles downtown) and hemmed in by amenities such as Forest Park, the site of the world's fair in 1904 which persuaded the rest of the United States that St Louis is not some outlying suburb of Chicago (more than 300 miles away). The setting is also a physical constraint on growth, but nobody wants that any more. The snag is that even successful departments may be smaller than a critical intellectual mass which is why, whatever the consequences of autonomy for arts and sciences, a measure of concentration may be prudent.

John Maddox

European geoscience union thrives

THE European Union of Geosciences (EUG), at its third meeting in Strasbourg on 1-4 April, appears to have established itself. Since EUG's foundation five years ago under C.J. Allègre of the University of Paris, attendance at the biennial meetings has increased from 600 in 1981 to more than 1,500 this year. The annual spring and fall meetings of the American Geophysical Union attract between 2,000 and 2,500 participants.

The need for such a forum in Europe was one of the stimuli for the creation of EUG, says the new president, Professor E.R. Oxburgh (University of Cambridge). There is a particular need, he says, for European earth science research students to meet and to present their work to more established people. To this end, costs are kept as low as possible. This year, more than 200 travel grants, valued at £10,000, were awarded for the meeting. In his presidential address, H. Wanke (Max

Planck Institute, Mainz) said that he knew of no comparable organization that spends so much of its income in this way.

EUG works on the principle that membership is strictly individual. Other than a part-time secretary provided by the European Science Foundation, the union has no permanent administration. Allègre says that EUG is essentially anti-bureaucratic.

Meetings are organized in turn by members from one country. The 1985 meeting was arranged primarily by the University of Cambridge, and the next (in 1987) will be organized by Eidgenössische Technische Hochschule, Zurich. After that, Italy will probably take over. All the meetings will take place in Strasbourg. Although some officers feel that this rotation of responsibility is a strength, others fear that growth may make a more permanent administration inevitable.

Peter Gambles

Hannover trade fair

Something for everybody

Hannover

THE world's largest and glossiest trade fair has become a supreme test of endurance. While the organizers boast of having crammed 60,000 cars into the car parks, yawning salesmen excuse themselves by saying they have just spent three hours in traffic.

Inevitably, innovation tends in these circumstances to be deeply buried in the products that people want to sell, preferably on the spot. But, after a week's spring sunshine, solar cells made of amorphous silicon abound, making enough electricity to circulate water in a toy swimming-pool or to light mock traffic lights. People like Interatom, a subsidiary of the nuclear energy constructor Kraft-Werke Union, say the opportunities for special uses are now rapidly increasing, the uneconomic character of solar cells notwithstanding.

Hannover's research ghetto, a building with the subtlety and floor-area of an aircraft hangar, is lavishly sprinkled with displays by West German universities of all kinds (of which there are three), most of them seeking industrial backing for research begun with public funds in academic laboratories. The phenomenon is partly instinctive and partly federal and local (*Land*) government policy. Industrial exploitation has become the watchword. (Some say it always has been.)

Academic stallholders hope to benefit in several ways. Professor Bert Küppers from the Fachhochschule at Aachen (a technical college not to be confused with the university at Aachen, called the Technische Hochschule) has a model of a generator capable of taking the variable-speed output from a windmill and producing constant-voltage 50 Hz current; if he is able to find a backer before the week is out, he will be in business.

Others look for industrial support to help win government grants. Professor Gerhard Müller from the Gesamthochschule at Essen, for example has developed a technique for coating objects made of materials such as plastic with metals of almost any kind. The principle is to use materials evaporated from an anode both as the working gas for a discharge and the coating substance. The development has cost the university less than DM 50,000 so far, but if the fair can recruit enough industrial testimonials, the chance of winning a big government grant will be bright.

On the biological front, pollution control and tree death are the prominent issues. The University of Oldenburgh has constructed a prototype of a self-contained system for processing sludge from sewage works which yields an oil-substitute rich in aromatic hydrocarbons and runs on energy by burning waste gases, so that the return of waste to the environment is kept to a

minimum. (The one per cent of residue contains tightly-bound metals, which may yet be reclaimed.) This is a project supported so far by the government which now needs industrial support.

The Karlsruhe nuclear research centre is on a similar tack with a pilot plant for burning rubbish in a self-contained environment. Heat makes electricity, and waste acid gases in solution can be used in reclaiming heavy metals from the solid waste.

Tree death is West Germany's anxiety just now, with the lumber industry an important part of the agricultural economy. The University of Freiburg and the Karlsruhe centre (the Kernforschungszentrum) have both developed techniques for identifying damaged trees from infrared photographs, and the use of these methods has already helped to suggest the sources of damaging pollution on the Western edge of the Black Forest. And people are growing trees in controlled test chambers, as the Gesellschaft für Strahlen und Umweltforschung in Munich is eager to proclaim, so that the physiological and biochemical effects of the various pollutants can be investigated individually.

This is the plentiful small change of the research on display, which is most of all impressive for the diversity and ingenuity of the academics. At the other end of the scale, the major research establishments under the umbrella of the Arbeitsgemeinschaft der Grossforschungseinrichtungen (big science spelled out) are also eager that passers-by shall recognize the useful things that they can do.

Radio-frequency resonators built as means of feeding power into particle accelerators (as at DESY, the Hamburg accelerator laboratory) are offered as being capable of other jobs. The Darmstadt laboratory (which accelerates heavy ions) also offers its beams as a way of punching micro-holes in pieces of material, if a little unconvincingly.

This group of a dozen national laboratories is exceedingly diverse and a little uncertain about its future. In the last year of the five-year government demand that spending should be reduced by 1.5 per cent a year, the organization has been cheered by the federal government's decision (see page 662) that Darmstadt should expand. But there is now some dismay that the umbrella organization should have been told, a few weeks ago, that it will have to make reductions of spending by differentiating among the twelve establishments, which range from the cancer research centre at Heidelberg to the national aerospace organization which is scattered over West Germany. People fear this new policy as a recipe for internecine strife in what was previously a self-contained organization.

The federal minister for research and technology, Dr Heinz Riesenhuber, gave away nothing on the degree of economy expected when he popped in by helicopter last week. (Helicopters land in the car park, whence ministers are ferried between cars full of machine-gun-toting police.) The minister shook hands or patted heads as appropriate, and held a public long-distance conversation of conspicuous banality with the West German research station in the Antarctic before being whisked away.

The Antarctic programme is nevertheless big — the annual budget runs at DM 65 million a year. One of its more intriguing sets of observations is that of the tides on the ice shelf at the entrance to the Weddell Sea, from which it is hoped eventually to disentangle simple ocean tides and those affecting the floating ice.

Industrial research is prominent, although electronics research is curiously inconspicuous. (Nobody has a stand devoted to the making of silicon chips.) The West German love affair with the motor car is illustrated by the large crowds around the Mercedes-Benz and Volkswagen stands. Preussac, the minerals extraction company, has a vivid exhibition of ores dragged up from the sea-bed, but is not especially hopeful that it will make money from its interest.

For the rest, the fair's technology is a curious contrast between the machines with moving parts and those which seem to have none. (The halls housing the mechanical machines are, by comparison, as quiet as cathedrals.) This, no doubt, is one reason why the organizers think they can safely split the fair next year. Most machinery-makers show something very old — Thyssen, for example, has a coal-burning locomotive outside its stand.

The computer people, by contrast, are all bustle. Apple, the US computer maker, has the best crowd-puller — it has simply provided 60 machines with which people can play if they fight their way to the head of the queue. Toshiba's illusionist is also a draw, but is there merely to suggest that his sponsors' photocopiers are magical. Given the time and energy, a person could decide which system will best meet his needs, although the computer-watchers are chiefly interested in computer-aided design and the like, which goes to show that this fair has become an essential part of West Germany's system of technical education.

Not everything is on this high plane. The lighting-hangar does give some manufacturers of specialized lighting systems a chance to show off, but it is mostly a display of lampshades of the kind that one might find (but would prefer not to find) in one's living-room. And an Australian company is selling (like "hot cakes") audio tapes of the sound of the ocean waves in which, it is said, there are hidden messages that will subliminally allow one to become better-adjusted in one of a dozen specific ways. How to live with traffic jams might be a good title for next year. □

Embryo research (contd)

SIR — The status of the early human embryo has been background music for much dancing on the head of a pin. The latest example is provided by C.B. Goodhart (*Nature* 14 March, p.126). He finds justification for "use of IVF embryos for experimental work expected to result in death" in the following argument. The fertilized egg *ex utero* is like an unfertilized egg in that both require external action to realize their potential to become an adult, the fertilized egg requiring activation by a sperm (or a substitute). But, the argument continues, fertilization does not really change "right to life" status because that right relates to being human and alive — which is characteristic of both gametes and zygotes. *Ergo*, since zygotes and early embryos *in vitro* will, like unfertilized eggs *in vitro*, die without external help, both are ethically available for experimentation unto death.

After vertigo subsides, one asks why the discovery does not apply to zygotes or early embryos in normal reproduction. Easy — "the embryo *in utero* has of itself (*sic*) the full potential to complete its development without any further intervention from outside". Possible parallel — the bank robber, once inside the bank and having cowed the cashier, has earned his ill-gotten gain because he can (barring police interference) carry it outside without further external aid.

But now, in truth, the embryos (blastocysts) *in utero* and *ex utero* have exactly the same potential, as demonstrated by normal development of *ex utero* blastocysts when transferred to a physiologically prepared uterus. Moreover, the state of both is very different from that of an unfertilized egg by virtue of the *genetic* individuality conferred by gamete fusion.

If there are good reasons why IVF embryos should be available for clinical investigation (and I believe there are), the justification is better provided through contemporary ethics and science, rather than from a return to scholasticism.

CLIFFORD GROBSTEIN

*Program in Science, Technology
and Public Affairs,
University of California, San Diego,
La Jolla, California 92093, USA*

SIR — C.B. Goodhart's argument (*Nature* 14 March, p. 126) that embryos fertilized *ex utero* lack independent viability and therefore are not potential human beings like their brothers and sisters in the womb after natural conception, is at first glance attractive, in that it might allow one for instance to be anti-abortion yet pro-experimentation on very early embryos; but it involves the debatable assumption that what comes naturally is outside our control in the sense that we can interfere with it only negatively, whereas for an embryo ferti-

lized *in vitro*, positive intervention is required for its survival. Is this a legally or morally valid distinction, especially for physicians and surgeons dedicated to interference with nature and rightly regarded as culpable if they do not do so when asked in the supposed interests of a patient, and might not feminists enquire whether their sex must regard themselves as the passive and vegetative receptacles of an embryo implanted in their womb by any male (or even by the Holy Ghost) who succeeds in depositing in them the necessary seed? Surely it is time we left the suspect word "natural" to advertisers of semisynthetic foodstuffs and took it upon ourselves to live up to the very worrying responsibility that we have for ordering our own destinies; or do we no longer assume that what man does is no longer part of nature since he is no longer mainly concerned with adaptation to his environment, but can adapt it to himself?

JOHN A. DAVIS

*Department of Paediatrics,
University of Cambridge Clinical School,
Level 8, Addenbrooke's Hospital,
Hills Road, Cambridge CB2 2QQ, UK*

Value-laden science

SIR — M. Hammerton's letter (*Nature* 31 January, p.343) demonstrates unwittingly that it is more difficult to identify value-free science than the unwary might realize.

The statement which Hammerton believes to be value free, "The lethal dose of cyanide is x gm/kilo body weight", contains implicit value judgements in the definition of "lethal dose" and how this relates to the real world, for example a coroner's inquiry. Is the lethal dose for a particular human individual determined better on the basis of extrapolation of LD₅₀s from animal experiments, or from a smaller sample of mortality data collected from other humans? Furthermore, independent of the definition of "lethal dose", if the forensic evidence indicates that the victim has a body burden of 0.2x gm/kilo, was cyanide the cause of death? Whatever the body burden of cyanide, what other factors ignored by the forensic evidence (such as the psychological state of the victim) could be relevant to the death? To answer such questions, value judgements are unavoidable.

Ironically, the statement which Hammerton believes is value-laden, "I am going to poison my spouse", can be interpreted as a statement of pure intention or as a prediction of one's own future behaviour. With these interpretations, although the statement is unpleasant, it contains no ethical, social, political or other evaluative judgements, and so can be considered to be value free. To make it clearly value-laden, one could re-interpret it as "Rather than (say) shoot my spouse, I

shall use poison", or one could modify it to "It is good/socially desirable/profitable to poison my spouse".

Although the matter is controversial amongst philosophers of science, people might wish to hold a "pure" mathematical statement, such as $2 + 2 = 4$, is value free in its content. However, even then the context could in some cases make it value laden: for instance, if it represented a reductionist approach to a scientific problem which some researchers believe can only be treated adequately by a holistic approach.

MARK DIESENDOFF

*CSIRO,
Division of Mathematics & Statistics,
Canberra ACT 2601,
Australia*

Sizewell inquiry

SIR — Your leading article (28 February, p.723) on the prospects for UK energy policy following the ending of the Sizewell inquiry and the coal industry strike raises a number of important questions. Unfortunately, it would demand too much space adequately to respond to all of them. I wish therefore simply to point out one factual error rather than to take issue on matters of interpretation.

You state that part of the explanation for the extreme length of the Sizewell inquiry (340 sitting days taking just over 2 years) was the inspector's decision at the outset to allow documentary evidence to be accepted only if it was read aloud, but this is not true. The inspector did require the proponent party, that is, Central Electricity Generating Board witnesses, to read their proofs of evidence in full so that their case would be fully set out in public. This is not so with the opposition. Groups or individuals who appeared in person were requested to summarize their evidence, sometimes to the considerable disadvantage of the objectors, as additional personal time had to be devoted to preparing such a summary in a cogent but no less coherent form.

In addition there were 112 entirely written submissions, of which the inspectors stated on the final inquiry day (p.79 in transcript): "I would like to add that I will take into account *all* these written representation in preparing my report."

I fear that because of your editorial stance that the structure and scope of the proceedings were inappropriate to the issues that ought in your judgement to have been under review, you may have overlooked some very important aspects of changes in energy policy that undoubtedly have emerged from the inquiry itself, such as nuclear waste management. No doubt these issues will persist in new, rather less exotic forums.

DAVID LOWRY

*Energy Research Group,
Open University,
Walton Hall,
Milton Keynes MK7 6AA, UK*

New ways with aggregation

An improved calculation with the classical equations of aggregation suggests new tasks for those who carry out computer simulations of these problems.

THE fashion for the study of aggregation in all its forms (from the formation of smoke particles to the polymerization of chemicals) is readily explicable. Quite apart from the several practical applications in which a better understanding would be valuable, computer simulations of aggregation processes have opened up an intriguing line of enquiry, confirming among other things that many aggregating particles do indeed have fractal, or non-integral, dimension.

General interest in aggregation is not, of course, new, but goes back at least to Smoluchowski's work in the 1920s. But how to make a bridge between the recent numerical work and the now-classical but phenomenological equations for which Smoluchowski is best known? Two people from the University of Utrecht, P.G.J. van Dongen and M.H. Ernst, have now made a brave attempt (*Phys. Rev. Lett.* **54**, 1396; 1985) which, among other things, has the virtue of suggesting where the computer people should next look.

The two approaches are as different as they could be. The computer simulations are typified by the model of T.A. Witten, which has been much elaborated since it first appeared in 1981. Particles are allowed to diffuse from the edges of some piece of lattice (square or 5 cubic, say), with instructions that they should stick to an aggregate grown from a single particle at the centre if they should happen to reach an immediately neighbouring position. Physically, the circumstances may be likened to those in which crystal growth is limited by the speed with which solute molecules can diffuse through a solution. Aggregates formed in this way do indeed have fractal dimensions. Recent elaborations of this simple model have allowed particles to shift their position after reaching the growth surface of an aggregate, or restrict occupancy of relatively inaccessible sites.

The Smoluchowski argument is more formal, and based on an enumeration of the ways in which aggregates of different sizes can stick together to form still larger aggregates. (The opposite process, disintegration, is supposed not to happen.) Aggregates containing N elementary particles will be formed (to a first approximation) only from pairs of smaller aggregates whose combined size is N . The rate at which this happens will be proportional to the number (or concentration) of each possible pair of constituent sub-aggregates,

that is to the product $c_j c_k$ where j and k add up to N . The snag is that the constant of proportionality will depend on the size of each member of the pair of sub-aggregates, so that the rate of formation of particles of size N is something like $K(j,k)c_j c_k$ summed over all possible pairs of constituent sub-aggregates. (A factor of two enters if the sum is written as a sum over both indices j and k .)

The rate at which aggregates of size N disappear is similarly calculated, but now aggregation with a particle of any size will remove an aggregate from the N -sized cohort. The result of adding these two rates together algebraically is a simple-looking set of equations expressing the rate of change of the quantities c_j which are nevertheless intractable (because they involve products of the variables, and are thus non-linear). But this, of course, is irrelevant because nothing is known about the coefficients $K(j,k)$. The hope that the computer simulations might throw some light on the nature of this function is one of the reasons why the subject has recently been fashionable.

Experiment, in the nature of things, can be of only limited help. Light-scattering from an aerosol or a colloid can, for example, provide information about the mean size of aggregating particles and some information about the size distribution. But the chance of defining the infinite set of coefficients $K(j,k)$ by the results of even the most refined experiment is plainly small. So the trick is to try to pin down the coefficients K , considered as a function of two particle sizes, from what scraps of information there may be. And a few generalizations are possible. In aggregating systems, it appears, the mean size does increase steadily with time. Experiments have also shown that in many kinds of aggregating systems, the evolving size-distribution is independent of the initial distribution when a sufficiently long time has passed.

Von Dongen and Ernst set out to generalize an earlier discussion of the problem (in 1966) by S.K. Friedlander. They restrict themselves to functions K which assume the form $j^p k^q$ when the aggregates are large, which is one sign of how much more needs to be done in this field. There are reasons for supposing that the sum of the two constants p and q cannot exceed 2.

Some of the results are nevertheless simple and striking. The formation of a gel from, say, a coagulating colloid requires that the sum of p and q should be greater

than unity. Indeed, it turns out that this condition is also sufficient in the sense that if the criterion is satisfied, the system will inevitably gel. And there is much that can be said about the size distribution of particles in a gelling system, both before and after they have set.

The chief interest of the new calculations will however turn on what they have to say about the scaling behaviour of the size-distribution function in all kinds of aggregating systems. Experiments as well as computer simulations suggest that the size distribution of particles in an aggregating system has a degree of uniformity in the sense that the variables c_j (which change with time) can be related to the average aggregation size (say s) by an expression of the form $s^{-r} f(j/s)$ where r is a constant and f some function of a single variable. Von Dongen and Ernst are able to calculate more about the characteristics of such a function, and of its implications, than has previously been possible.

That there should be a simplification of this kind is really remarkable. What it implies is that, with the passage of time, an aggregating system will evolve in such a way that the size-distribution of the particles is similarly related to the mean particle size. Part of what van Dongen and Ernst succeed in doing is to calculate the properties of the unchanging distribution function, in practice by calculating its moments (the mean, variance and so on).

While the results obtained in this way are not inconsistent with what has been learned from computer simulation, there is clearly a great deal to be learned about the relationship between the simulations and the mathematical form of the coagulation function. One would have thought that there is now plenty of promise in attempts to simulate the aggregation process for a few of the special circumstances in which van Dongen and Ernst have shown the equations to be reasonably tractable.

A particularly valuable goal would be the definition of the exact form of the coagulation function when most aggregates are still small. The obvious difficulty is that the numerical programs so far used in these simulations are devoid of much in the way of the physics of what happens when substantial aggregates stick together. To what extent does rearrangement then become possible, for example? That is also a challenge for experimenters. May it be one of the unplanned benefits of the interest in nuclear winter?

John Maddox

Sex chromosomes and evolution

Haldane's Rule OK

from J.S. Jones and Nick Barton

THERE are few rules in biology, and most of those which have been proposed — Cope's Law of the Unspecialized, Haeckel's Biogenetic Law and many others — have sunk into obscurity or have been so confused by exceptions as to lose much of their force. One survivor of more than sixty years is Haldane's Rule: "When in the F_1 progeny of two animal races one sex is absent, rare or sterile, that sex is the heterozygous (heterogametic) sex"¹. From the paper of J.A. Coyne on page 736 of this issue the genetic basis of this rule can now be seen to reflect some fundamental processes in developmental biology, models of sex determination and speciation².

Haldane's rule applies not only in the familiar cases in which it is the males which are heterogametic — for example, the sterile XY male donkeys which result from a cross between horse and an ass — but also in those animals (such as birds, butterflies and snakes) in which the female is the heterogametic sex. Hence males are not innately more sensitive to the effects of interspecific hybridization than are females. Haldane himself (and many of his successor's) felt that the sterility of heterogametic hybrids results from an imbalance between sex chromosomes and autosomes: XX hybrids have two complete haploid sets while XY hybrids lack the full complement of chromosomes from one parental species that is necessary for fertility. Coyne, however, has now shown that, in *Drosophila* at least, this is not true, and that a different mechanism is involved.

Drosophila simulans can hybridize with *D. mauritiana* and *D. sechellia*. Making use of an attached-X stock of *D. simulans*, Coyne generates hybrid females in which both X chromosomes come from *D. simulans*, but one set of autosomes comes from each of the parental species. Such flies have the same degree of discordance between X chromosomes and autosomes as do XY hybrid males, and might be expected to be sterile. Instead, they are quite fertile. X-autosome interactions cannot explain the sterility of heterogametic hybrids: Coyne suggests that it must arise from genetic interactions between the X and Y chromosomes of the parental species.

He confirms this suggestion by producing hybrid flies which (as far as is feasible given the genetic limitations of the stocks available) differ only in their Y chromosomes. Substitution of a *D. mauritiana* Y chromosome into flies which have a *D. simulans* X chromosome leads to a 96 per cent drop in the production of motile sperm, indicating that interactions between X and Y are indeed responsible for their hybrid sterility. Similar interactions might also be responsible for the failure of repro-

duction in species in which females are the heterogametic sex, as inter-racial crosses of the moth *Lymantria dispar*, in which inappropriate X and Y chromosomes are combined, fail to produce females³. However, such interactions between sex chromosomes cannot explain all cases of hybrid sterility. Male sterility occurs in interspecific crosses between some grasshoppers which have an XO sex-determining mechanism (and hence lack a Y chromosome)⁴.

It has long been thought that the sex chromosomes diverged from a pair of homologues differing only at a locus determining sex. Meiotic pairing between X and Y in man is quite extensive⁵, suggesting that there is homology between the two. A 36-kilobase stretch of single-copy DNA which has been translocated to the human Y chromosome during recent hominid evolution has been shown to diverge by less than one per cent from an equivalent sequence on the X chromosome^{6,7} and a 30,000 M_r protein is coded for by a gene on the tip of the human X chromosome which is also present on the Y chromosome⁸.

Male sex determination in mammals demands the cooperation of both sex chromosomes⁹. The Y is necessary for testis and sperm formation: human XO individuals are phenotypically female, and in some marsupials that are mosaics for XO and XY cells, spermatogonia are always XY. There is a testis-determining locus on the mouse Y chromosome, mutants of which can lead to male sterility¹⁰. The X chromosome is also implicated in disorders of male sex determination. In the primary spermatocyte of normal males, the single X chromosome is inactivated. Any retention of its function interferes with fertility: XXY humans are sterile (perhaps because of the presence of an active X in spermatocytes), and translocation of part of the X to an autosome can also lead to male sterility. The Y chromosome may be involved in the repression of the X which is necessary for successful spermatogenesis¹¹; this may be why XX human males and 'sex-reversed' male XX mice (both of which have undergone a translocation of part of their Y chromosome to the X) show early degeneration of germ cells^{12,13}. Sterility of this kind may result from failure of the Y-linked repressor locus in its new chromosomal milieu; male mice in which the sex-reversed allele is on the Y chromosome are quite fertile.

Direct genetic evidence for the existence in *Drosophila* of Y-linked genes that suppress related loci on the X has recently emerged. The *D. melanogaster* Y carries several genes responsible for male fertility, most of which encode components of

sperm¹⁴. Deficiencies of any of these lead to a series of errors in spermatogenesis. One such mutant is known as *stellate*. Flies bearing this allele suffer from a derangement of meiosis and the production of star-shaped crystals in the primary spermatocyte¹⁵. A cloned segment of DNA which hybridizes with the *stellate* locus has now been found to contain a 1,250 base-pair segment repeated at least eight times in tandem, which is present on both the X and Y chromosomes¹⁶. In XO individuals (which in *Drosophila*, unlike mammals, are male) the X-linked copies of this region produce up to 70 times more RNA than is produced by the *stellate* locus in normal XY males, providing direct evidence that genes on the Y are repressing others with a related function on the X, and further support for the view that male fertility involves interaction between these chromosomes.

The distribution of the *stellate* sequence shows considerable differences among various members of the *D. simulans* group. Hence the sterility of heterogametic hybrids — Haldane's rule — may arise from a failure of X-chromosome repression by a foreign Y. The same seems to be true in mice¹⁷. When a Y chromosome from *Mus domesticus* is transferred into a laboratory strain of *Mus musculus*, XY individuals develop as infertile females or as hermaphrodites, probably because *M. domesticus* Y chromosomes bear a testis-determining locus which cannot interact with the appropriate X-linked locus in *M. musculus*. Although an autosomal sex-determining locus may also be necessary for male fertility in these hybrid mice¹⁸ there is a parallel with hybrid sterility in *Drosophila*.

This new evidence for the importance of the sex chromosomes in hybrid sterility raises some general questions. Why, in most groups, is the germ line more sensitive to hybridization than is the soma; and why should genes that lead to reproductive isolation so often be sex-linked? One interesting possibility might involve 'selfish' DNA, propagated through the germ line despite its deleterious effects¹⁹. Co-evolution between such DNA and genes that are able to suppress its replication will therefore be most intense in the germ line, which might therefore be unduly sensitive to the mobilization of selfish DNA on hybridization. Examples of selfish genes include B chromosomes in Orthoptera and, in *Drosophila*, segregation distortion, cytoplasmic sterility and the mobile genetic elements that cause hybrid dysgenesis²⁰⁻²³. These genes have their greatest effect in the germ line, and their replication may be suppressed by modifiers elsewhere in the genome. It may be more than coincidence that the *stellate* locus involves a tandem repeat, whose structure is similar to other repeated elements in the genome, and whose taxonomic distribution is similar to some of the mobile elements responsible for hybrid dysgenesis²⁴.

As well as the genes for hybrid sterility, a disproportionate number of interspecific

differences seem to be sex linked. Not enough information is available on protein or DNA structure to test whether X-linked structural loci show undue divergence among species²⁵. However, many genes for mating preferences are on the X chromosome in *Drosophila* and some butterflies^{26,27}, and sex chromosome rearrangements are frequently found when related Orthoptera are compared²⁸. In the Australian grasshopper *Caledia captiva*, the hybrid zone for the sex chromosomes is much narrower than are those for the autosomes; there are stronger linkage disequilibria for sex-linked characters; and there is much less introgression between races for characters borne on the X chromosome than for those on the autosomes²⁹. All this, together with the marked differences between male and female viability which emerge in reciprocal backcrosses, shows that the sex chromosomes are subject to much stronger selection pressures in hybrids than are the others.

This close involvement of sex chromosomes with species differentiation might arise in several ways. If speciation involves population bottlenecks (as is often suggested), then genetic drift will act more strongly on the relatively small population of X chromosomes than on the autosomes. In addition, sex-linked differences are subject to less selection against heterozygotes and recombinants in hybrids than are those carried on autosomes. Lande's estimates³⁰ of the rates of chromosomal evolution in insects and mammals imply an excess of sex-linked differences which is sufficient to account for those observed in nature. In addition, genes under sexual selection (which may be a potent force in the evolution of reproductive isolation³¹) may evolve more rapidly when sex linked, and the 'hitch-hiking' effect of closely linked loci may also be more important for genes

on the X chromosome^{25,32}.

Predictions of the mechanisms of DNA replication, of molecular mimicry by parasites and of kin selection and selfish genes can all be recognized in Haldane's early writings³³. His discovery of the regularities which underly the sterility of hybrids seems set to become equally important in modern biology. □

- Haldane, J.B.S. *J. Genet.* **12**, 101 (1922).
- Coyne, J.A. *Nature* **314**, 736 (1985).
- Clarke, C.A. & Ford, E.B. *Proc. R. Soc. B* **206**, 381 (1982).
- Butlin R.K. & Hewitt G.M. *Biol. J. Linn. Soc.* (in the press).
- Chandley, A.C. *et al. Cytogenet. cell. Genet.* **38**, 241 (1984).
- Page, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5352 (1982).
- Page, D., Harper M.E., Love, J., & Botstein, D. *Nature* **311**, 119 (1984).
- Goodfellow, P. *et al. Nature* **302**, 346 (1983).
- Gordon, F.W. & Ruddle, F.H. *Science* **211**, 1265 (1981).
- Eicher, E.M. & Washburn L.L. *J. exp. Zool.* **228**, 297 (1983).
- Gueilaen, G. *et al. Nature* **307**, 172 (1984).
- Evans, E.P., Burtenshaw, M.D. & Lindsley, D.L. *Nature* **300**, 443 (1982).
- Burgoyne, P. *Nature* **307**, 109 (1984).
- Goldstein, L.S.B., Hardy, R.W. & Lindsay, S.L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7405 (1982).
- Hardy, R.W. *et al. Genetics* **107**, 591 (1984).
- Livak, K.J. *Genetics* **107**, 611 (1984).
- Eicher, E.M. Washburn, L.L., Whitney, J.B. & Morrow K.E. *Science* **217**, 535 (1982).
- Washburn, L.L. & Eicher, E.M. *J. exp. Zool.* **228**, 297 (1983).
- Doolittle, W.F. & Sapienza, C. *Nature* **284**, 601 (1980).
- Shaw, M.W. *Biol. J. Linn. Soc.* **23**, 77 (1984).
- Hartl, D. *Lect. Not. Biomath.* **19**, 65 (1977).
- Ehrman, L. in *Coevolution* (ed. Futuyma, D.J. & Slatkin, M.) 128 (Sinauer Press, Sunderland MA, 1983).
- Engels, W.R. *A. Rev. Genet.* **17** (1982).
- Hartl, D.L. Hiraizum, Y. in *The Genetics and Biology of Drosophila* Vol. 1b (eds Ashburner, M. & Novitski, E.) (Academic, New York, 1976).
- Avery, P.J. *Genet. Res.* **44**, 321 (1985).
- Zouros, E. *Genetics* **97**, 703 (1981).
- Gruha, J.W. & Taylor, O.R. *Evolution* **34**, 673 (1980).
- Hewitt, G.M. *Animal Cytogenetics: Orthoptera* (Gebrüder Borntraeger, Berlin 1979).
- Moran, C. *Heredity* **42**, 13 (1979).
- Lande, R. *Evolution* **33**, 234 (1979).
- Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3721 (1981).
- Maynard Smith, J. & Haigh, J. *Genet. Res.* **23**, 23 (1974).
- Haldane, J.B.S. *On Being the Right Size* (ed. Maynard Smith, J.) (Oxford University Press, 1985).

J.S. Jones and Nick Barton are in the Department of Genetics and Biometry, University College London, London NW1 2HE, UK.

Chemistry

Activating the C-H bond

from J.M. Thomas

THE paper by Ito and Lunsford on page 721 of this issue¹ should interest chemists of all complexions, and others too, who are aroused by the challenge of effecting controlled activation of C-H bonds, especially in molecules such as methane (CH₄) the simplest of saturated hydrocarbons and the main constituent of natural gas.

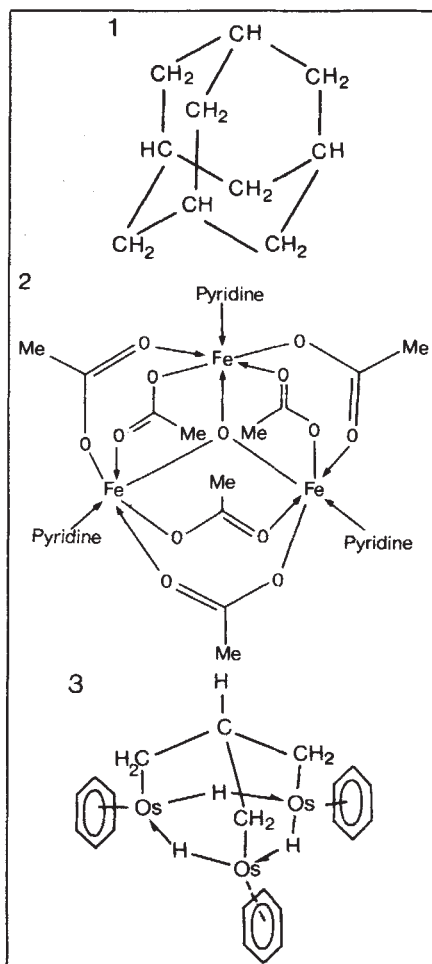
It is not difficult to sever all four C-H bonds simultaneously: the ease with which this complete and relatively uncontrolled oxidation is accomplished is what makes natural gas a fuel. Ways of partially oxidizing CH₄ to produce synthetic gas ('syngas') are also well known: this is the first step in present-day industrial methods of catalytically synthesizing methanol, acetic acid and many other products, including vinyl and acetate polymers². Much

scope exists for exploiting syngas in new ways. The integrated process, now operated on a pilot scale by Haldor Topsøe in Copenhagen, for synthesizing gasoline using a zeolite catalyst is one example; the laboratory synthesis by Jones *et al.*³ of acetylene using a de-intercalated graphite catalyst is another. But how far have we progressed in rivaling the comparative ease with which nature performs subtle, milder oxidation of relatively stable molecules?

For a start, organic⁴ and organometallic⁵⁻⁸ chemists have recently been devising ways of emulating the performance of enzymatic systems, such as the non-haem iron protein ω -hydroxylase or the iron-porphyrin-based cytochrome P-450 which selectively hydroxylate unactivated C-H bonds. And theoretical chemists have

been trying to identify the factors that influence the transfer of electrons between a transition metal atom or surface and the bonding and antibonding orbitals of the C-H bond in simple paraffins⁹.

In addition, at least two distinct lines of experimental attack have been used to activate C-H bonds in saturated hydrocarbons by chemical means. Organic chemists have tended to focus on molecules rich in tertiary and secondary C-H bonds, such as adamantane (1). These bonds are a little easier to activate than the primary bonds of methane. Encouraging results have been achieved using a tri-iron cluster compound (2) which, in the presence of other reagents, catalyses the conversion of adamantane to corresponding alcohols and adamantanone at room temperature⁴. Like a variety of metal surfaces, bare metal atoms, under the right conditions, are undoubtedly capable of activating the C-H bond. For example, compound (3) is prepared by the co-condensation of osmium atoms and a



1:1 mixture of benzene and 2-methyl propane⁸. Its precise mode of formation is unknown, but presumably involves intermediate steps in which unsaturated metal centres of discrete osmium intermediates activate the C-H bonds of 2-methyl propane. The very existence of (3) contributes to our general understanding of how a saturated hydrocarbon may be adsorbed on to the surface of a catalyst.

In the other approach, favoured by many

physical chemists, ways have been sought of achieving the direct partial oxidation of methane to methanol, ethylene and benzene. To date, such work has entailed the use of very high temperatures, and many of the more successful experiments have used nitrous oxide as oxidizing agent, which could never figure as conveniently as O₂ or air in any commercially viable process. In that respect, the use of O₂ as oxidant, reported by Ito and Lunsford in this issue¹, is a significant advance in the general quest to oxidize methane selectively. Moreover, the catalyst they use is devoid of transition metal ions, which have been thought to be necessary constituents of efficient catalyst formulations. The mechanism of activation at the active site (believed to be a Li⁺O⁻ pair in the lithium-doped magnesia catalyst) probably entails the formation of methyl radicals.

Gas-phase kineticists and organic chemists have, over the years, recognized the mechanistic and preparative importance of free radicals in numerous categories of reactions. But for several decades past, the role of surface-generated free radicals in

heterogeneous catalytic reactions, at least in the types of reaction studied by Ito and Lunsford, has been regarded as minor. Both the preparation of solids that facilitate the formation of reactive free radicals when hydrocarbon molecules impinge on their surfaces and the role of surface-generated free radicals will doubtless receive lively future attention in formulating new strategies for the mild oxidation of methane. □

1. Ito, T. & Lunsford, J.N. *Nature* **314**, 721 (1985).
2. Falbe, J. *Phil. Trans. R. Soc. A* **300**, 205 (1981).
3. Jones, W., Schlögl, R. & Thomas, J.M. *J. chem. Soc. Chem. Commun.* 464 (1984).
4. Barton, D.H.R., Gastinger, M.J. & Motherwell, W.B. *J. chem. Soc. Chem. Commun.* 731 (1983).
5. Crabtree, R.H., Mihelcic, J.M. & Quirk, J.M. *J. Am. chem. Soc.* **101**, 7738 (1979).
6. Watson, P.L. & Parshall, G.W. *Acc. Chem. Res.* **18**, 51 (1985).
7. Nizova, G.V., Krevor, J.V.Z., Kitaigorodskii, A.I. & Nizova, G.D. *Izv. Akad. Nauk SSSR* **12**, 2805 (1982).
8. Green, M.H.L. & O'Hare, D. *J. chem. Soc. Chem. Commun.* 355 (1985).
9. Saillard, J.-Y. & Hoffmann, R. *J. Am. chem. Soc.* **106**, 2006 (1984).

J.M. Thomas is in the Department of Physical Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EP, UK.

Developmental biology

Homoeo boxes and strings for the packaging of genes?

from Brigid Hogan

NINETEEN eighty-four was a vintage year for developmental biology. The discovery of the 'homoeo box', first as a common feature of several developmental genes in *Drosophila*, and then as a highly conserved sequence in vertebrate DNA, has catalysed a lively fermentation of ideas from disparate areas of research. One outcome has been the isolation of mouse and human genes containing homoeo-box sequences. Now, on page 713 of this issue¹, Colberg-Poley *et al.* provide the first evidence that homoeo-box sequences are transcribed in mammalian cells: the fact that this transcription seems to peak transiently during the differentiation of mouse embryonal carcinoma cells is bound to change some preconceived ideas about the function of mammalian homoeo-box genes, and opens up many new possibilities for experimentation.

As readers who have followed the story so far will know, the *Drosophila* homoeobox is a DNA sequence of about 180 nucleotides, discovered in at least five genes in the Antennapedia complex, three in the Bithorax complex and five outside these major gene clusters^{2,3}. Mutations in specific homoeo-box containing genes can disrupt the normal pattern of segmentation either by causing a switch in segment identity (for example, from one carrying antennae to one carrying legs, in the case of *Antennapedia*) or by changing their number or anterior-posterior polarity.

There is good evidence in *Drosophila* that homoeo-box sequences encode domains of proteins that are located in the nucleus and are possibly involved in DNA binding^{4,5}. The general consensus, then, is that homoeo-box genes in *Drosophila* act in different combinations to select specific sets of genes defining unique properties of each of the compartments making up the larval segments. At least eight sequences closely related to the *Drosophila* homoeo box have been found in the mouse and human genome. In the case of the mouse gene described by Colberg-Poley *et al.*, there is a 95 per cent match between the amino-acid sequence that its homoeo box is predicted to encode and that of the homoeo box of the *Antennapedia* gene of *Drosophila*. Comparison of *Drosophila* and mammalian genes also points to similarities in, for example, gene clustering, 3' intron/exon organization and the protein sequences beyond the strict confines of the homoeo box.

All this provides a strong impetus to search for a function for the mammalian homoeo-box genes. One approach is to test whether they are identical to any of the genes that, when mutated, cause mouse or human developmental defects. Already we know that one mouse gene is on chromosome 6 (the *Hox-1* locus)^{6,7}, and two groups^{7,8} have recently assigned other homoeo-box genes to mouse chromosome 11 (*Hox-2*), and to the long arm of human

chromosome 17 (*Hu-1* and *Hu-2*), near the loci for thymidine kinase, galactokinase and pro- $\alpha 1$ (I) collagen genes. These three genes have been mapped to chromosome 11 in the mouse (ref. 8; K. Harbers and R. Jaenisch, personal communication) and it is reasonable to assume that *Hox-2* is close by. This brings within the range of speculation⁸ the possibility that the gene at the *Hox-2* locus is identical to *Tail-short* (*Ts*), a gene that is lethal when mutant and homozygous, and maps close to thymidine kinase and galactokinase on mouse chromosome 11.

Fortunately for today's molecular biologists, many mouse developmental mutants have been the subject of much classical embryological and genetic study. Reference to the work of Deol⁹ on the embryology of mice with one normal and one mutant *Ts* gene shows that although they develop many interesting skeletal abnormalities, early defects are also seen in non-segmented tissues, including a marked reduction in the number of blood islands in the extra-embryonic yolk sac. More recently, Patterson has described how mouse embryos that are homozygous for *Ts* die before implantation and can be detected as abnormal even by the eight-cell stage¹⁰. So, if there is allelism between *Ts* and *Hox-2* (or another homoeo-box gene), at least one homoeo-box gene product is expressed at a very early stage and possibly in a wide range of tissues. Such speculation is not without foundation, as a homoeo-box gene of *Xenopus* is transcribed in the oocyte, and then again from the neurala stage at least up to the swimming tadpole¹¹.

Detailed studies on the transcription of homoeo-box genes in the mouse embryo will depend on good *in situ* hybridization data. Meanwhile, Colberg-Poley *et al.* have looked at transcription in F9 mouse embryonal carcinoma cells in culture. There are many tumour and embryo-derived cell lines now available, with a range of developmental potentials — from F9 cells, which hardly change at all unless treated with retinoic acid, to EK or ESC cells, which are difficult to restrain from differentiating into diverse cell types. Even when treated with retinoic acid, F9 cells give rise only to cells that express markers of extra-embryonic tissues, either visceral or parietal endoderm depending on the exact conditions¹². In differentiating along these alternate pathways, the cells transcribe new sets of genes and switch off others. Colberg-Poley *et al.* find that a new 2.4-kilobase RNA, which hybridizes to their homoeo-box probe, appears in F9 cells 17–24 h after the induction of differentiation; four days later this RNA is no longer present. The authors hint that RNA which hybridizes to the same probe is also present in normal mouse embryos, in both embryonic tissues and extra-embryonic tissues.

These results suggest that the pattern of mammalian homoeo-box gene transcript-

ion may be very complex and is not necessarily confined to specific regions of the embryo during the time when the serially repeated somite-derived tissues are being established, as might be expected by analogy with their expression in *Drosophila* embryos. Perhaps homoeo-box genes are needed for a much wider range of decision-making processes in mammalian embryos — unless, of course, they have nothing to do with development at all. For exploring these possibilities, F9 cells can be grown in large amounts so that cDNA libraries can easily be produced, and antibodies to fusion proteins used to try and isolate homoeo-box-containing proteins.

Such proteins promise to have extremely interesting properties. A possible homology between a 20-amino-acid subdomain of the homoeo box and the *MATa1* and *MATa2* proteins of yeast, and the helix-2-turn-helix-3 region of known DNA-binding proteins such as bacteriophage λ Cro protein and *Escherichia coli* CAP protein, has been emphasized^{4,13,14}. However, the rules governing prediction of three-dimensional structure from amino-acid sequence are not yet sufficiently refined for these homologies to be more than suggestive. Another motif of homoeo-box proteins is the presence of long strings of a single amino acid in regions flanking the homoeo box; repeats

of glutamic acid (15 at the 3' end of the clone described in ref. 1) glutamine, alanine and proline have been reported^{1,4,15}. These strings may act as spacers between different domains of the protein designed for functions such as DNA binding, RNA binding or interaction with other regulatory proteins. Perhaps assemblies of boxes and strings act in various combinations to pack up DNA that is no longer needed, or to link new batches of active genes to the nuclear matrix. □

1. Colberg-Poley, A.M., Voss, S.D., Chowdhury, K. & Gruss, P. *Nature* **314**, 713 (1985).
2. Gehring, W.J. *Cell* **40**, 3 (1985).
3. Scott, M.P. *Trends Genet.* **1**, 74 (1985).
4. Laughon, A. & Scott, M.P. *Nature* **310**, 25 (1984).
5. Beachy, P.A., Helfand, S.L. & Hogness, D.S. *Nature* **313**, 545 (1985).
6. McGinnis, W., Harl, C.P., Gehring, W.J. & Ruddle, F.H. *Cell* **38**, 675 (1984).
7. Rabin, M. *et al.* *Nature* **314**, 175 (1985).
8. Joyner, A.L. *et al.* *Nature* **314**, 173 (1985).
9. Deol, M.S. *Proc. R. Soc. B155*, 78 (1961).
10. Patterson, H.F. *J. exp. Zool.* **211**, 247 (1980).
11. Muller, M.M., Carrasco, A.E. & De Robertis, E.M. *Cell* **39**, 157 (1984).
12. Hogan, B.L.M., Barlow, D.P. & Tilly, R. *Cancer Surv.* **2**, 155 (1983).
13. Pabo, C.O. & Sauer, R.T. *A. Rev. Biochem.* **53**, 293 (1984).
14. Miller, A.M., MacKay, V.L. & Nasmyth, K.A. *Nature* **314**, 589 (1985).
15. Poole, S.J., Kauer, L.M., Drees, B. & Kornberg T. *Cell* **40**, 37 (1985).

Brigid Hogan is at the National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

Archaeology

Megalithic recycling in Brittany

From Paul G. Bahn

A REMARKABLE discovery made and recently published by C.-T. Le Roux¹ links two of the most famous megalithic monuments in Brittany — the tomb of Gavrinis, known for the superbly carved orthostats lining the passage, and the passage grave of La Table des Marchands at Locmariaquer, which also bears fine examples of megalithic art².

The underside of the capstone of the Table des Marchands (lower section of figure) has deep carvings of a beautiful hafted axe with two loops, a 'crook' and a couple of 'commas'. In addition, at the broken end, there are more lines which some observers have thought to be a "curious motif with a double crook"², whereas others saw them as part of a quadruped. Because part of this engraving is obscured by an orthostat, it has seemed likely that the slab formerly stood upright with its decoration clearly visible, but was then re-used. This hypothesis has also been made for the capstone of the tomb of Mané Rutual, which also has carving on its underside and is almost 12 m in length³.

Le Roux's discovery has finally brought proof of this hypothesis. During restoration work on a capstone at the tomb of Gavrinis (now located on an island that was still linked to the mainland when these monuments were constructed), carvings

were revealed on the upper side that had been invisible ever since the tomb's tumulus was set in place. The slab had clearly been dressed by man at the sides, and was broken at either end. Its carvings (central section of figure) comprised a great 'hache-charrue' (axe-plough), most of a horned animal and the horns of a second. Despite some resemblance to a mouflon or ibex, it is clear from the thick muzzle, short neck, straight back, angular rump and double-curved horns that a bovid is depicted.

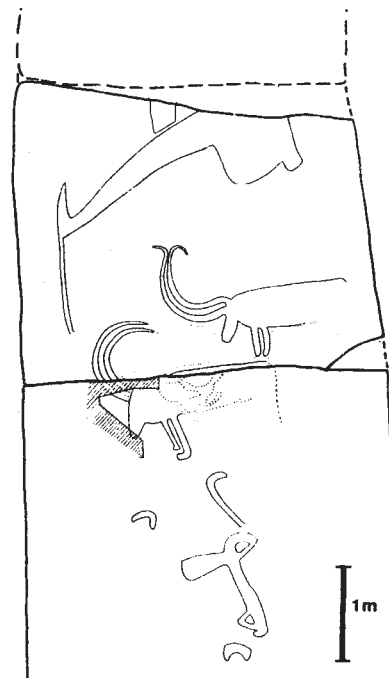
The dimensions and grain of the rock, the shape of its break and the nature of its motifs all led Le Roux to link it with the Table des Marchands capstone. Sure enough, the two fragments would fit together, revealing a second bovid. Moreover, a similar rock used for a capstone in the long tumulus of 'er-Vinglé', just north of the Table des Marchands, was probably broken off from the other end of the Gavrinis slab (upper section of figure).

If so, the reconstituted stela would have stood about 14 m high, 3.7 m wide and 0.8 m thick, weighed about 100 tons and been decorated on the central portion of one face. It would have been second in size only to the Grand Menhir Brisé, 'er-Grah', which now lies in four fragments beside the Table des Marchands, and was over 20 m in length. It too was deliberately dressed

and has an 'hache-charrue' carved on one of its fragments².

Because two of the three fragments of the 'new' stela are at Locmariaquer, it is likely that the block originally stood somewhere nearby, perhaps in the company of the Grand Menhir Brisé and the great carved endstone of the Table des Marchands. This means that the Gavrinis slab, of about 7.5 m³ and weighing 20 tons, was transported 4 km (as the crow flies) across several water courses.

It has been suggested that the discovery implies a megalithic phase that preceded the tombs and involved great decorated stelae which, later, were deliberately thrown down and broken with wedges into fragments used in the building of tombs⁴; as the carvings were hidden to view at Gavrinis, they must have lost all significance. Le Roux is more cautious and envisages coexistence and complementarity of the menhirs and the tombs. Dating is not good enough to distinguish between the two possibilities. Gavrinis was perhaps built around 5,200–5,000 BP, and its use may have ended about 4,500 BP; but we have no idea when the stelae were put up or how



The decorated slabs of Gavrinis (above) and Table des Marchands (below), as they would have fitted together (after C.-T. Le Roux).

long they lasted. Their decoration may provide proof of a very early use of the plough and of bovids for traction in the western-most part of Europe. But many questions about the history and function of these impressive but enigmatic monuments remain unanswered. □

1. Le Roux, C.-T. *Bull. Soc. Préhist. fr.* **81**, 240 (1984).
2. Shee Twohig, E. *The Megalithic Art of Western Europe* (Clarendon, Oxford, 1981).
3. L'Hélégouach, J. in *Bretagne: Livret-Guide de l'Excursion* (ed. Giot, P.-R.) 129 (IXe Congr. U.I.S.P.P., Nice, 1976).
4. Mohen, J.-P. *La Recherche* **15**, 1528 (1984).

Paul G. Bahn can be contacted via the School of Archaeology and Oriental Studies, University of Liverpool, Liverpool L69 3BX, UK.

Astronomy

A giant black hole at the centre of a dwarf galaxy

from C. Martin Gaskell

THE dwarf elliptical galaxy M32, one of the satellite galaxies of the nearby spiral galaxy M31, has recently been shown to have a central dark object with a mass five million times greater than the Sun¹. This is the most convincing case yet of a massive black hole in the centre of a galaxy.

Signs of violent activity are to be found in the nuclei of many galaxies. They range from modest goings-on in our own Galaxy to the tremendous outpourings of energy in quasars. The source of this energy is believed to be the accretion of matter onto massive black holes^{2,3} but the evidence is still only indirect. One obvious way to detect a massive black hole is through its gravitational attraction on nearby stars. In 1978, the late Peter J. Young and his collaborators^{4,5} claimed to have found just such an effect; in the very centre of the giant elliptical galaxy M87, the average speed of the stars shoots up dramatically, as though by sudden strong gravitational pull. M87 is certainly a galaxy in which one would expect a massive black hole because it is relatively active, but the claim soon came under attack. Theoretical studies⁶⁻⁸ showed that the observations of Young and his collaborators^{4,5} could be explained without invoking the presence of a black hole (of 5×10^9 solar mass) if the orbits of the stars were predominantly radial (rather than random). Further observations⁹ showed a dense concentration of stars in the centre of M87, which also had to be taken into consideration. The presence of a black hole in M87 remains possible but is certainly not necessary.

M32 has perhaps a thousand times fewer stars than a giant elliptical galaxy like M87 and is certainly not an active galaxy. It does, however, show an unexplained concentration of stars towards its centre. Using the 200-inch telescope on Mt Palomar to study the motions of the stars in and near this central concentration, John Tonry has now found that, as with M87, the dispersion in the velocities of the stars rises suddenly towards the centre. But unlike M87, the system of stars as a whole is clearly rotating in M32, so radial orbits

cannot be the explanation; there has to be some excess mass there. Moreover, this mass is more than can readily be accounted for by the stars alone; an extra five million times the mass of our Sun is present, according to Tonry. A massive black hole is the most likely explanation.

If there is a black hole of such size in the middle of M32, why is the galaxy totally inactive? The answer is probably that there

is nothing to fuel the black hole; gas is escaping from the galaxy rather than falling into the centre, as it would in a more massive galaxy.

M32 is by no means unique in having a dense central condensation of stars. It just happens to be very near and is therefore easy to study. To make comparable observations on more distant examples will need better resolution than can be achieved from the ground. When the Space Telescope (due to be launched towards the end of next year) turns on the nuclei of these galaxies I, for one, am pretty confident that they too will show signs of massive objects lurking in their centres. □

C.M. Gaskell is in the Astronomy Department, University of Texas, Austin, Texas 78712, USA.

Vision

Fireworks in the retina

from David I. Vaney

VISUAL information undergoes a sophisticated coding process in the retina, culminating in the diverse and functionally complex output of the ganglion cells to the brain: The receptive fields of the ganglion cells are shaped by the preceding interneurons, which include a set of neurons without axons, termed amacrine cells. The responses of amacrine cells are poorly documented because it is difficult to record from their small cell bodies, and the responses obtained may not reflect the local interactions between adjacent input and output synapses on the amacrine branches. In a recent series of elegant papers, however, R.H. Masland and his colleagues elucidate the functional properties of one small population of amacrine cells, those that contain acetylcholine¹. Their findings suggest that the plexus formed by the branches of these cells provides the substrate for motion detection within subunits of a ganglion cell's receptive field.

In the rabbit retina, about four per cent of the amacrine cells synthesize and release acetylcholine. Their cell bodies are symmetrically distributed on both margins of the inner plexiform layer (IPL, the narrow zone containing processes of ganglion cells and presynaptic interneurons) and their terminals form two dense bands within the IPL^{4,5}. These cells are of particular interest to retinal physiologists because they are implicated in the generation of direction selectivity in two types of ganglion cells⁶. Masland *et al.* have neatly side-stepped the problems of electrophysiological recording by studying the morphological characteristics and pharmacological responses of the whole population of these cells, using a sophisticated range of techniques. Although the shape and distribution of these neurons were presaged by previous studies⁷⁻¹⁰ which used classical histological methods, this integrated approach provides

definitive descriptions and fresh insights.

Masland, Mills and Hayden¹ have shown that the acetylcholine-synthesizing amacrine cells identified by autoradiography are also selectively labelled *in vivo* by the fluorescent dye diaminodiphenylindole (DAPI), so that the cells can be identified and mapped in a whole-mounted retina. Tauchi and Masland² have impaled the DAPI-labelled cells with a micropipette under visual control and iontophoretically injected the fluorescent dye Lucifer yellow into the cytoplasm. This reveals a characteristic dendritic morphology that has been compared to a 'starburst' firework³, and confirms the previous description by Vaney using similar methods¹¹ (Fig.1). The arborization ranges from 200-800 μm in diameter but is virtually planar, with the distal dendrites narrowly stratified in one of the two bands in the plexiform layer that contain acetylcholine.

Masland, Mills and Cassidy³ have demonstrated that the two morphological subtypes of cholinergic amacrine cells respond differently to visual stimuli; those branching in the inner half (on the ganglion-cell side) of the plexiform layer release a transient burst of acetylcholine at light ON, whereas those with branches in the outer half of the layer release a burst at light OFF (Fig.2). This has been established convincingly by monitoring the distribution and release of labelled acetylcholine in response to visual stimulation *in vitro*, while using a pharmacological agent known to block the transmission of ON responses.

The most striking feature of the cholinergic amacrine cells is the extensive overlap of their dendrites in the plexiform layer: each point on the retina is covered by the dendritic trees of 25-70 cells of each subtype^{2,11}, an order of magnitude greater than the overlap of the horizontal or ganglion cell arborizations in the cat retina¹².

1. Tonry J.L. *Astrophys. J. Lett.* **283**, L27 (1984).
2. Zeldovich, Y.B. & Novikov, I.D. *Dokl. Acad. Nauk. SSSR* **158**, 811 (1964).
3. Salpeter, E.E. *Astrophys. J.* **140**, 796 (1964).
4. Young, P.J., Westpal, J.A., Kristian, J. & Wilson, C.P. *Astrophys. J.* **221**, 721 (1978).
5. Sargent, W.L.W., Young, P.J., Boksenberg, A., Shortridge, K., Lynds, C.R. & Hardwick, F.D.A. *Astrophys. J.* **221**, 731 (1978).
6. Duncan, M.J. & Wheeler, J.C. *Astrophys. J. Lett.* **237**, L27 (1980).
7. Binney, J. & Mamon, G.A. *Mon. Not. R. astr. Soc.* **200**, 361 (1982).
8. Tonry, J.L. *Astrophys. J.* **266**, 58 (1983).
9. Dressler, A. *Astrophys. J. Lett.* **240**, L11 (1980).

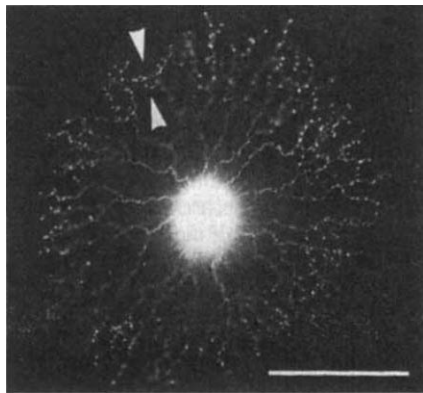


Fig. 1 Cholinergic amacrine cells in rabbit retina filled with the fluorescent dye, Lucifer yellow; the arborization in the inner plexiform layer is sharply focused beneath the blurred cell body. The characteristic appearance of these interneurons resembles a 'starburst' firework⁹, an effect created by the numerous varicosities on the fine branches in the distal annular zone (between arrows). Scale, 100 μ m. (Micrograph by D.I. Vaney).

Nearly all the output synapses of cholinergic amacrine cells are found on the varicose branches in the distal zone of the arborization, and make contact with ganglion cell dendrites¹⁰. Although the 50–80 terminal dendrites of a cholinergic amacrine are closely spaced in the plexus, a group of distal terminals may actually be separated from an adjacent group by a thin dendritic branch 200–800 μ m long. Groups of terminals may thus be electrically isolated from each other and so form multiple subunits within a single arborization¹². Tauchi and Masland² propose that the functional unit of cholinergic action is an aggregate of such subunits from overlapping dendritic trees. The transient local release of acetylcholine from these restricted regions of the plexus, at light ON and OFF, could provide the temporal and spatial resolution required for motion detection within the receptive field of single ganglion cells (Fig. 2).

The cholinergic plexus is so dense that there is a dendritic varicosity every few microns and, theoretically, the spatial resolution of the system could be as fine as the separation between varicosities. In practice, however, the resolution of the cholinergic plexus is limited by the less dense mosaic of presynaptic bipolar cells,

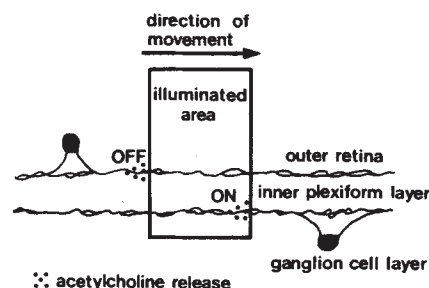


Fig. 2 Local release of acetylcholine from amacrine cell terminals in response to a moving stimulus (modified from ref. 3).

which makes the overlap of branches appear curiously redundant. Moreover, dendritic replication seems a cumbersome way of producing transmitter amplification "capable of powerful local action"². Such arguments assume that the cholinergic amacrine cells are functionally equivalent, but this need not be the case³: one scheme for producing direction-selective (DS) receptive fields in ganglion cells requires that the cholinergic subunits are anisotropic⁶.

DS ganglion cells respond with vigorous excitation when a stimulus is moved through the receptive field in the preferred direction, whereas motion in the reverse ('null') direction inhibits them, evoking little or no response¹³. The ON-OFF DS units in the rabbit retina form four groups, each with a range of preferred directions that does not overlap the other groups. Cholinergic amacrine cells provide substantial direct input to DS ganglion cells and, if potentiated by the cholinergic agonist physostigmine, their excitation can overcome the null-direction inhibition⁶. It is not known, however, whether the nonlinear interactions between the excitatory and inhibitory inputs take place in the ganglion cell itself or in the presynaptic neurones. In the latter case, the release of acetylcholine would itself be directionally

selective. DS ganglion cells in one group may thus receive cholinergic input from different subunits to those in the other three groups, and this would further rationalize the great overlap of the cholinergic amacrine cells.

The apparent simplicity of directional selectivity has seduced retinal physiologists for over twenty years but its mechanism still eludes them. The advances made by Masland and his colleagues provide new tools for dissecting the phenomenon and suggest that satisfaction is now within reach. □

1. Masland, R.H., Mills, J.W. & Hayden, S.A. *Proc. R. Soc. B* **233**, 79 (1984).
2. Tauchi, M. & Masland, R.H. *Proc. R. Soc. B* **223**, 101 (1984).
3. Masland, R.H., Mills, J.W. & Cassidy, C. *Proc. R. Soc. B* **223**, 121 (1984).
4. Masland, R.H. & Mills, J.W. *J. Cell Biol.* **83**, 159 (1979).
5. Hayden, S.A., Masland, R.H. & Mills, J.W. *Science* **210**, 435 (1980).
6. Ariel, M. & Daw, N.W. *J. Physiol., Lond.* **324**, 161 (1982).
7. Hughes, A. & Vaney, D.I. *J. comp. Neurol.* **189**, 169 (1980).
8. Vaney, D.I., Peichl, L. & Boycott, B.B. *J. comp. Neurol.* **199**, 373 (1981).
9. Famiglietti, E.V. *Brain Res.* **261**, 138 (1983).
10. Famiglietti, E.V. *Vision Res.* **23**, 1265 (1983).
11. Vaney, D.I. *Proc. R. Soc. B* **220**, 501 (1984).
12. Miller, R.F. & Bloomfield, S.A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3069 (1983).
13. Barlow, H.B. & Levick, W.R. *J. Physiol., Lond.* **178**, 477 (1965).

David I. Vaney is at the National Vision Research Institute of Australia, 386 Cardigan Street, Carlton, Victoria 3053, Australia.

Tectonics

Whence British Columbia?

from E. Irving

IDEAS about the origin of the western cordillera of North America are undergoing an evolution that resembles, in miniature, that which occurred in global tectonics about 30 years ago, when mobility replaced fixity as the basis for geological thought. The palaeomagnetic measurement of very large displacements among continents in the mid-1950s was followed by a decade of scepticism, until evidence from the oceans made the case for global mobility unassailable. Similarly, measurements in the early 1950s of very large displacements in deformed belts within the confines of a modern continent met a large measure of disbelief; the orthodox view was that mountain belts were formed more or less where they are now. In the North American context, attitudes soon changed as geologists came to realize that the western cordillera was made up of a patchwork of dissimilar blocks which came together in the mid-Mesozoic (150–100 Myr ago) from sundry and not accurately known places. The notion of mobility therefore began to be accepted in a loose qualitative sense.

Recently, geologically young displacements have been found further and further eastward in the cordillera, especially in the Canadian sector, as if the assembled patchwork of blocks had moved northwards as a whole minicontinent during the Laramide Orogeny (80–50 Myr ago), when

the Rockies formed. This turn of events has not been received with enthusiasm by geologists because they see little or no evidence for it in the rock record. However, it is possible that much of the direct evidence for this massive event has been obscured by deep blind thrusts, as yet unrecognized, along which large-scale horizontal transport has occurred. Rather than investigate the suggested movement of the western minicontinent, most workers have focused on easier problems concerned with its earlier assembly — with local rather than regional geology. Chamberlain and Lambert in their paper on page 707 of this issue¹ step back from the local scene and try to describe the Laramide event in its entirety.

Their scheme is based on two premises. First, the assembled minicontinent (dubbed by them 'Cordillera', although I would have preferred the more accurate name 'Baja British Columbia' following Stone's² terminology) was situated much further south than at present, as indicated by palaeomagnetism^{3,4}. Second, Cordillera was driven northwards because it was attached to an oceanic plate², the Kula Plate⁵.

Within this general kinematic framework Chamberlain and Lambert envisage two stages. To begin with, there was an interval of closure as Cordillera moved from somewhere offshore in the Pacific Ocean towards mainland North America, a small oceanic

basin being consumed between them. Geological evidence for this lost ocean is to be found, according to Chamberlain and Lambert, in granitoid rocks formed when the oceanic crust melted as it sank beneath Cordillera. They argue that the volcanoes that formerly surmounted these granitoid bodies were responsible for the first appearance of airborne volcanic ash in the Alberta Basin to the east about 100 Myr ago. I cannot comment on how well their proposal agrees with the geological evidence, but it does not agree with palaeomagnetism, according to which Cordillera, at that time, was as much as 2,000 km south of Alberta⁶. Moreover, it could not have moved so early if, as Chamberlain and Lambert argue, it moved with the Kula Plate, because the latter did not form (by splitting from the Farallon Plate) until 85 Myr ago⁷. Cordillera actually began its journey after Chamberlain and Lambert suppose it to have arrived; this does not disprove their scheme, but means only that the timing needs revision.

During the second envisaged stage, Cordillera has reached mainland North America and proceeds to slide sideways (strike-slip motion) along the Rocky Mountain Trench. It is this trench that Chamberlain and Lambert choose as the present-day boundary between Cordillera and North America. The trench is a linear series of narrow valleys situated west of the Rocky and Mackenzie Mountains and stretching some 2,000 km from the Alaskan border, through Canada and south to the continental United States.

There is evidence of up to 1,000 km of motion along or close to the northern part of the trench⁸ but across the southern trench there are similarities in Phanerozoic

rocks that have usually been taken to preclude strike-slip motion. Chamberlain and Lambert regard these similarities as superficial and the consequence of wholesale Laramide thrusting across the trench, which has obscured the deeper break. In their search for this break, they point to the Proterozoic Malton gneiss complex, located astride the southern trench. They claim to have demonstrated geochemical and age differences between rocks of the complex to the east and west of the trench, and argue that the trench is a window of disrupted basement rocks exposed through overlying Laramide thrust-sheets. The earlier interpretation — that the malton gneiss 'locks' the trench, denying motion along it — is neatly turned on its head, so that it becomes evidence for motion rather than against it. Geochemists will want to see the analyses and structural geologists will wish to reassess the disposition of these remote rocky outcrops, but the argument is arresting, and not lightly set aside.

I am glad that Chamberlain and Lambert have grasped this nettle. It is not difficult to pick holes in their scheme (it could hardly be otherwise with so complex a problem) but they are certainly looking at the problem in the correct way. Theirs is a useful and timely contribution which will assist discussion of the nature of the Laramide Orogeny. □

1. Chamberlain, V.E. & Lambert, R. *St J. Nature* **314**, 707 (1985).
2. Stone, D.B. *Proc. 6th A. Symp. Geol. Soc. Alaska*, 1 (1977).
3. Beck, M.E. & Nosen, L. *Nature* **235**, 11 (1972).
4. Monger, J.W.H. & Irving, E. *Nature* **285**, 289 (1980).
5. Page, B.M. & Engenbreitson, D.C. *Tectonics* **3**, 133 (1984).
6. Irving, E., Woodsworth, G., Wynne, P.J. & A. Morrison, A. *Can. J. Earth Sci.* (in the press).
7. Woods M.T. & Davies, G.F. *Earth planet. Sci. Lett.* **38**, 161 (1982).
8. Gabrielse, H. *Bull. geol. Soc. Am.* **96**, 1 (1985).

E. Irving is at the Pacific Geoscience Centre, Box 6000, Sidney, B.C., V8L 4B2, Canada.

Genetic screening

The heredity of haemophilia

from Miranda Robertson

GENE therapy is not yet a reality. But the technology that may one day make possible the molecular repair of genetic defects is already providing a range of diagnostic and screening tools that must complicate the task of genetic counselling even as they widen the choice open to the parents of a potentially defective child. The easiest kind of genetic disease to detect by molecular means is one in which the defective gene is known and the rate of spontaneous mutation is low. This is the case for some of the hereditary anaemias, for some of which it is now possible to make gene probes that will distinguish a mutant from a normal globin gene even where the disease is due to a single point mutation. The case of haemophilia, in which the defective gene is also known and has recently been isolated and cloned, is more complicated, as a report from Gitschier *et al.* on page 738 of this issue clearly shows¹.

The commonest form of disease is haemophilia A, which affects about one individual in 10,000. The defect is in the gene encoding factor VIII, one of the cascade of plasma proteins that interact in blood clotting: haemophiliacs die from internal haemorrhage unless treated with exogenous factor VIII. The gene for factor VIII is on the X chromosome, so usually only males are affected, females being protected by the normal gene on their second X chromosome. Females may thus be symptomless carriers: a woman whose father is haemophiliac is of necessity a carrier of the gene; one whose brother or uncle is haemophiliac may be. Treatment with factor VIII from donor plasma is effective but carries the risk of hepatitis and more recently — but very rarely — of acquired immune deficiency syndrome (AIDS); moreover about 15 per cent of haemophiliacs develop antibodies to the replacement factor VIII.

In the circumstances many female carriers would rather have an abortion than an affected child. Factor VIII, however, cannot be assayed in an unborn child until 18–20 weeks into pregnancy, and in the absence of any way of distinguishing the mutant from the innocent maternal X chromosome, some carriers have opted for early abortion of all male fetuses, accepting the 50 per cent chance that a normal child will be destroyed.

The announcement late last year that the gene for factor VIII had been isolated, characterized and cloned^{2,3} opened up the possibility of precise identification of the gene defect and, with it, reliable detection of carriers and accurate prenatal diagnosis within the first 8–11 weeks of pregnancy, when abortion is safer and less traumatic. As it turns out, however, the factor VIII gene is proving less cooperative than might have been hoped.

Ideally, a mutant gene is distinguished from a normal one by a radiolabelled DNA probe that detects the mutation itself. In practice, this is really only applicable where a few identified mutations have spread through a population and probes can be tailored to each of them; where a high incidence of genetic disease is due to a high rate of spontaneous mutation, the molecular defect may be different for each affected family. This now seems to be the case for haemophilia A⁴. Moreover, as a point mutation can disable the factor VIII gene as effectively as a major deletion⁴, relatively gross screening, which would only pick up a truncated gene, will miss many defects whose consequences are severe.

The alternative to direct detection of the mutation is to identify the mutant gene in individual families by means of natural sequence variations (polymorphisms) present in the population as a whole. Such polymorphisms have no effect on the function of the gene, and simply serve as markers by which one factor VIII gene can be distinguished from another. Gitschier *et al.* have discovered such a polymorphism in the factor VIII gene, and have been able to use it in carrier detection.

As a screening tool, detection of polymorphisms has several grave disadvantages by comparison with direct detection of mutations. First, because any given polymorphism may be present in perfectly normal genes as well as the mutant gene, accurate diagnosis means having enough pedigree information to establish which polymorphic variant of the gene carries the haemophilia mutation in a given family. Given adequate pedigree information, the value of a polymorphism will then depend on two more factors. The first is the number of polymorphic variants — usually there are only two; and the second is their frequency in the population at large. If there are only two variants (say, I and II) and one of them (say II) is infrequent, then a particular carrier is quite likely to be homozygous — that is, both her factor VIII genes are likely to be variant I, although

only one of them will contain the mutation; the polymorphism will then be uninformative. (In the case of autosomal defects — where the mutation is not on the sex chromosomes — the position is further complicated by the possibility that the father may also carry the polymorphism in a normal gene; but this problem does not arise in X-linked diseases such as haemophilia because a boy inherits only one X chromosome, from his mother.)

Polymorphisms are most easily detected by bacterial enzymes (restriction enzymes) that cut DNA at specific sequences. A variation in the target sequence will prevent the enzyme from cutting at that point, and change the sizes of the fragments into which the gene is cut by the enzyme. Gitschier *et al.* experimented with 37 restriction enzymes, and found only one polymorphism in the 90,000 base pairs of the factor VIII gene that they searched¹. This is, as they point out, remarkably low compared with, for example, globin genes. The polymorphism that they did find is revealed by the restriction enzyme *BclI*, and has two alleles, present at roughly 70 and 30 per cent in the population. This makes it reasonably informative: five of eight tested carriers were heterozygous.

Much more informative, however, is the polymorphism recently reported by Oberle *et al.*⁵. It lies outside the factor VIII gene itself but is in the DNA that flanks the gene and so is usually inherited with it. Termed St14, the site has 10 polymorphic variants.

The disadvantage of this marker, as with any linked marker, is that genetic recombination can occasionally separate the marker from the gene giving a false result in a screening test. Nevertheless, St14 and *BclI* combined with another linked marker, DX13, provide a powerful genetic screening tool in those families where the necessary pedigree data are available.

Even with routine screening, given the high rate of *de novo* mutation and the existence of a normal carrier state, haemophilia A is not going to disappear from human populations. There is much talk of using genetically manipulated bone marrow to correct genetic defects where, as in haemophilia, the site of action of the gene is the plasma. This would entail the introduction of a normal factor VIII gene into the patient's own bone-marrow cells, which do not normally express the gene. At this stage, such therapy could not be justified in the face of unknown risks, for a disease treatable by more conventional, albeit imperfect, means. Nor is it clear whether the problem of antibody production against factor VIII could be avoided, for instance by very early bone-marrow therapy. In all likelihood, we shall see. □

1. Gitschier, J., Drayna, D., Tuddenham, E.G.D., White, R.L. & Lawn, R.M. *Nature* **314**, 738 (1985).
2. Wood, W.I. *et al.* *Nature* **312**, 326 (1984).
3. Gitschier, J. *et al.* *Nature* **312**, 330 (1984).
4. Gitschier, J. *et al.* *Nature* (in the press).
5. Oberle, I. *et al.* *New Engl. J. Med.* **312**, 682 (1985).

Miranda Robertson is Biology Editor of *Nature*.

Mathematics

Feigenbaum's fixed function

from Ian Stewart

FOR several years there has been a flurry of activity on the topic of 'deterministic chaos' whereby quite simple systems, described by equations having no explicit random terms, nevertheless exhibit behaviour that seems to be stochastic. Using ideas from renormalization group theory (originally invented to handle problems in quantum field theory), Mitchell Feigenbaum of Cornell University was led to conjecture some very specific 'universal' properties of such systems. Feigenbaum's conjecture was proved in full rigour by M. Campanino, H. Epstein and D. Ruelle (*Commun. math. Phys.* **79**, 261; 1981 and *Topology* **21**, 125; 1982) and by Oscar Lanford (*Bull. Am. math. Soc.* **6**; 1982). Both proofs made heavy use of complicated (though exact) numerical estimates, which required computer assistance. Remarkably, Lanford — who works at the Institut des Hautes Etudes Scientifiques in Bures-sur-Yvette — has now improved the proof so that only a calculator is required.

Feigenbaum's conjecture arose from a numerical study of the behaviour of the nonlinear recurrence equation $x_{n+1} = kx_n(1-x_n)$, as the constant k increases from 0 to 4. An initially stable fixed point bifurcates into a pair of points of period 2, which in turn split into periods 4, 8, 16, ... and so on. By the time k reaches a value around 3.7, an infinite number of these bifurcations has occurred; this is called a period-doubling cascade. Such cascades are of interest in a variety of applications, notably fluid turbulence. Feigenbaum was interested in the rate at which successive steps in the cascade developed. Let k_n be the value of k at which the n th period-doubling occurs. The ratio of gaps between consecutive bifurcation points is $\delta_n = (k_{n+1} - k_n)/(k_{n+2} - k_{n+1})$; as n tends to infinity the value of δ approaches 4.6692016... That is, the successive bifurcation points accumulate according to a geometric progression at a rate measured by δ . On repeating the calculation for other nonlinear recurrence equations, Feigenbaum discovered that the values of δ were identical. This coincidence has important implications for experimental confirmation of mathematical models based on nonlinear recurrence equations (discrete dynamical systems), because by observing the values of k at which the first five or six period-doublings occur, one may obtain an experimental estimate for δ , to be compared with a very precise theoretical prediction.

Further investigation suggested a reason for the 'universality' of δ . If $f(x) = kx(1-x)$ is the function being iterated in the recurrence equation, then it turns out that the second iterate $f^2(x) = f(f(x))$ has almost the same shape of graph as f , pro-

vided the axes of the graph are similarly positioned and scaled, and provided attention is confined to a suitable range of values of x . The clue to the universality lies in finding a function g for which this scaling process is exact. In other words, g must satisfy the Feigenbaum-Cvitanovic functional equation

$$g(x) = -g(g(-ax))/a$$

for a positive constant a (representing the required scale factor). The difficulty is to show that such a function g exists.

To do so, one starts with a function that satisfies this equation to a very high degree of approximation and uses it to construct a sequence of functions, by a process of renormalization, that converges to the desired solution. Abstractly one thinks of the desired g as living in a space X of functions: the renormalization operator T transforms any function h in X into a new function Th . The operator T is chosen so that the Feigenbaum-Cvitanovic equation boils down to the requirement that g be a fixed point (or fixed function) of T . Standard theorems, concerning fixed points of operators on function spaces, then finish the proof.

The difficulty is to obtain a sufficiently close initial estimate. It is here that the computer comes in. Although in the normal run of things the numbers held in a computer are subject to rounding error, it is possible to keep track of the largest size such errors can be, and hence keep exact estimates of the degree of accuracy of the calculations. In the present case, the initial choice of function h is somewhat complicated and so, therefore, are these estimates.

Lanford has succeeded in simplifying the procedure to the point at which the estimates may be performed using a hand-held calculator (here a Hewlett-Packard HP-15C, accurate to nine decimal places). The proof remains delicate, but is much simpler than its predecessors. Note, however, that to apply g to period-doubling, some additional properties must be established, which can still only be obtained using Lanford's computer-assisted proof.

The essence of the problem seems to be that the Feigenbaum fixed function g cannot be related in any very direct way to more standard functions, such as polynomials, for which methods of calculation can easily be found. Functional equations themselves have been something of a mathematical backwater. The prospect of applying them to nonlinear dynamics now makes it important to develop new and better techniques to handle their solutions. In such a way may new mathematics be born from a good physical problem. □

Ian Stewart is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Mushroom cloud the result of a meteor?

SIR — On 9 April 1984, a number of airline pilots witnessed a mushroom cloud above the northern Pacific Ocean, about 400 kilometres off the east coast of Japan. The cloud was first witnessed by the pilots of Japan Airlines flight JAL 036 at 14:06 GMT (23:06 local time) while flying route A-90 eastbound near the checkpoint PAWES. The location of the first sighting was 38.5° N, 146.0° E. The pilot reported that a large spherical cloud rose from the cloud deck below his aircraft and expanded rapidly to its full height of 60,000–70,000 feet and its full diameter of 200 miles in something between 1 min 50 s and 2 min 30 s. There was a heavy deck of clouds at the 14,000 feet level, but visibility was fair because of moonlight. About 12 s later, near 42° N, 151° E, they felt light turbulence. JAL 036 dispatched a May-day call at 14:21 GMT, put the crew on oxygen, and descended. JAL 036 was diverted to Elmendorf Air Force Base in Anchorage, Alaska, where radioactivity tests were conducted, but no trace was found. There was no communications breakup, nor any problems with the aircraft instruments.

Radioactivity tests conducted on other aircraft reporting the cloud also found no trace. Subsequent interviews with those pilots showed that the cloud exerted no fireballs or flash of lights. However, slight luminosity of the cloud was reported by several pilots. The pilot of KLM 868 reported that at first the cloud was opaque, then as it got larger it became possible to see stars through it. The phenomenon was visible for about 55 min.

The event was investigated by the US Department of Defense, Japanese Defence Agency, Federal Aviation Administration (FAA) and many others. All investigations concluded that it was not a nuclear explosion and it was possibly a natural phenomenon, but they failed to give a reasonable explanation.

We have investigated the event by querying FAA pilot interviews, newspaper reports, weather satellite photographs, and geophysical data from various institutions. Geophysical data include seismic, hydrophone, atmospheric electricity, telluric current, geomagnetic, and microbarograph data. However, none of the data we have studied contains a detectable signal from this event. Measurements of SO₂ in the ozone layer made by the National Aeronautics and Space Administration were also negative, suggesting that it was not a volcanic event.

We also obtained comment and opinions from experts in various related fields about this event. Some of the possible explanations we have pursued are:

- (1) A giant cumulus cloud.
- (2) A vapour plume from the Kaitoku Seamount, 1,500 km south-west of the event.
- (3) Burning of rocket fuel causing the

ionization of cloud particles and the subsequent formation of a plume.

- (4) A meteor causing the ionization of cloud particles and subsequent formation of a plume.

Of these possibilities, we have concluded that only the meteor plume has sufficient energy to explain the event. Meteors can enter the Earth's atmosphere at almost any angle. The speed of a meteor ranges from 15 to 110 km s⁻¹, and its height ranges from 150 km to ground level. Meteors slow down considerably upon entering the Earth's atmosphere. Normally, when a meteor does reach the ground as a meteorite, it travels the final portion of its trajectory as a dark body; it is rarely luminous below heights of 10 to 20 km.

A meteor's energy is very high. For example, the muzzle velocity of a 30-calibre bullet is about 840 m s⁻¹. A meteoroid the weight of a bullet (10 g), entering the lower atmosphere at a speed of 30 km s⁻¹ has as much kinetic energy as a bullet leaving a muzzle. At this speed, even cloud particles striking at the surface of the meteor can produce a shattering effect. Collisions with atmospheric particles ionize atoms and produce visible luminosity.

Our hypothesis is that a meteor encountered the cloud deck and almost simultaneously shattered into pieces. Most of the kinetic energy was converted to heat which evaporated cloud particles, and the hot gas formed a plume. Because of the high speed of the meteor, the shattering took place over a distance spanning a few tens of kilometres, thus forming a large plume. The large distance over which the shattering occurred explains why only a slight disturbance was felt by the pilot and why microbarographs in Japan failed to detect the signal. It explains why there was neither a fireball nor a flash of light associated with the event; and it also explains the slight luminosity of the mushroom cloud.

If our explanation is correct, the pilot's experience on the night of 9 April may be the first close encounter of an aircraft with a meteor. We would welcome readers' comments on this event.

ANDRE C. CHANG
JAMES A. BURNETT

*Teledyne Geotech,
PO Box 334,
Alexandria,
Virginia 22313, USA*

A mechanism for prion replication

SIR — The putative causative agents of some infectious diseases which attack the brain, such as Creutzfeldt-Jakob disease of man and scrapie of sheep, have been called 'prions'. Prions consist of protein, and apparently they contain no nucleic acid¹, although this contention has been challenged². As reviewed recently in *News and Views*³, two alternative mechanisms

have been proposed for the replication of prions. (1) Activation of transcription of a host gene coding for prion protein. (2) 'Reverse translation' or protein-directed protein synthesis. There is no direct evidence for either of these mechanisms and therefore the replication of prions remains an enigma, although possible mechanisms for self-replication of proteins have been proposed⁴.

We wish to propose an alternative possible mechanism for the replication of prions, which could operate if prions are indeed devoid of nucleic acid.

Prions consist of a single major protein component called⁵ PrP, of relative molecular mass 27,000–30,000. We propose that PrP is a very stable slow-acting proteolytic enzyme, which may produce brain pathology precisely because of its protease activity. We further propose that the mechanism by which PrP 'replicates' is proteolytic attack of a precursor protein (which we call pre-PrP) present in many cells, and certainly in the brain, having an as yet unknown normal function. This precursor is folded in a configuration that has no proteolytic activity but, being a 'substrate' of PrP, it can be proteolytically cleaved with production of a fragment (or domain) which is PrP itself. The newly released PrP is then able to digest other pre-PrP molecules, thus bringing about multiplication of an apparent 'infective agent'. This proteolytic multiplication is reminiscent of other well-known cases in which a proteolytic enzyme (such as pepsin) acts on a precursor (such as pepsinogen) splitting off a portion to yield more of the same active protease.

The model we propose is consistent with the 'autocatalytic' time course of the diseases in question, in which there is a long incubation period followed by rapid disease progression. Moreover, the prior existence of pre-PrP in the body would account for lack of immune response (characteristically observed in these diseases), and ought to be detected also in normal cells, either directly or through its messenger.

We hope experiments will be devised to test this hypothesis, which may have relevance to other so-called 'slow infections' and degenerative diseases of the brain.

MAURIZIO BRUNORI

*Istituto di Chimica,
Facoltà di Medicina,
Università di Roma "La Sapienza",
Pza A. Moro 5,
00185 Rome, Italy*

BERNARD TALBOT

*National Institute of Allergy
and Infectious Diseases,
Bethesda,
Maryland 20205, USA*

1. Prusiner, S.B. *Science* **216**, 136–144 (1982).
2. Rohwer, R.G. *Nature* **308**, 658–662 (1984).
3. Masters, C.M. *Nature* **314**, 15–16 (1985).
4. Griffith, J.S. *Nature* **215**, 1043–1044 (1967).
5. Prusiner, S.B. *et al.* *Cell* **38**, 127–134 (1984).

The ragged edge

Reflections upon what is to be found in books, besides the text.

DR JOHNSON said about cucumbers that they should be well sliced, and dressed with pepper and vinegar, and then thrown out as good for nothing. This is more or less what happens when scientific information of similar insipidity is processed into books, that are catalogued and at once fade, all but a few of them, from human consciousness. To browse among the review copies that each week fill some yards of shelf space in offices of journals such as this quickly induces a desolating sense of the futility of man's endeavour. Who will actually read this tome containing the *Proceedings of the Fifth Annual Conference on Isoelectric Focussing in Polyacrylamide Gels*? Will the narrative perhaps go with more of a swing than the sinking feeling in one's viscera portends? It will not. It will be what Lewis Carroll subtitled *The Hunting of the Snark* — an agony in five fits. Why do they do it then, those battalions of authors and editors? I think it was a historian who defined research as reading two books that have never before been read in order to write a thesis that will also never be read. Perhaps Byron hit the nail on the head:

*'Tis pleasant, sure, to see one's name in print;
A book's a book, although there's nothing in't.*

There is material here from which more than one PhD thesis could be hewn. If there is little pleasure or improvement to be had from the texts of most books, can we find it elsewhere between the covers? There is not much fun in the index, the contribution all too often, it seems, of the author's wife. The best one can hope for is a diverting typo or two: what for instance are these congenial metabolic diseases to which this one keeps referring? And what are the particular virtues of freeze-frying? But no, there are more abundant pickings for the sociologist of science nearer the title page. The dust-flap of course often yields a morsel for the connoisseur. Here is one that comes to me at second-hand, by way of the *New Yorker*; it commends the book, a dictionary, gift bound in tooled cowhide, and notes that it is "alphabetized for easy reference". Good idea, comments the *New Yorker*. But it is more likely that you will merely be assured that you hold in your hand an integrated wide-parameter, in-depth state-of-the-art overview, that will enable the student to digest this seminal cornerstone of a rich and pregnant field; it will set new benchmark standards as a very nearly unique quantum jump, an indispensable framework, comprehensive yet concise, directed as much at the specialist research worker as at the undergraduate, and essential reading for astronomers, cytologists, palaeontologists, cytopalaeoastonomers, palaeocytologists and astro-

palaeontologists. But since the credit for such gems of the blurb-writer's craft cannot be laid on the author, let us start at the beginning with the epigraph. This may be in Latin or even Greek (in which the author does not have to be fluent), but if in English it should preferably be totally inscrutable. Thus: "A philosopher does not need a torch to gather glow-worms by at mid-day" — Ernest Bramah.

There follows the preface with a list of acknowledgements in which wives and family feature by reason of their fortitude in sharing for several weeks the arduous of a scribe's life. More particularly, wives without exception have typed the manuscript, compiled the index, read the proofs or supplied candid and perceptive criticism. How many such wives are wives still by the time the book is in print? The late Queenie Leavis was said to have discovered only on publication that she was not co-author of one of her husband's blockbusters. There has been a rash of lawsuits in recent years, from which emerged sorry stories of help-meets, who not only filled fountain-pens, typed manuscripts and cooked meals, but also earned money to see the husband through his PhD and were then returned to store for their pains. The scientific wife is a social species ripe for a PhD of her own.

It is customary for the author also to render thanks to a colleague or two for critical reading of the text; it would be an almost indecent breach of tradition not to conclude with the arch avowal that what errors remain are his — but the implication is that they are his friends' just the same. (As Charles II said in response to Rochester's famous mock-epitaph, "He never said a foolish thing/Nor ever did a wise one" — his words were his own, but his deeds were his ministers). There is then much citing of sources from which material has been gouged. I have only once seen a trickle of bile, such as many authors must have suppressed with difficulty; it comes from Jonathan Gathorne-Hardy's splendid treatise on the British nanny: "I must also formally acknowledge the following publishers, institutions and people for their permission, often after substantial payment, to quote from the books whose copyright they control...". The sycophantic expressions of gratitude to patrons have alas largely vanished, for the realization has grown that institutions, even those as munificent as the National Institutes of Health, have no mechanism for receiving homage or converting it into gratification. To thank the NIH has always seemed to me a little like the action of the small boy who was observed by the Reverend Sydney Smith to be stroking the shell of a tortoise

in order, as he said, to please the creature; "why", said Sydney Smith, "you might as well stroke the dome of St. Paul's to please the Dean and Chapter". More entertainment was to be had from writers' relations with patrons of an earlier age. Dr Johnson's dissatisfaction with Lord Chesterfield's performance in the office is recorded in the *Dictionary*, in which he defines a patron as "one who countenances, supports or protects. Commonly a wretch who supports with insolence, and is paid with flattery". He enlarged on this in his celebrated letter to Chesterfield, in which he drew the parallel between the patron and one who observes the struggles of a drowning man and only when he reaches the shore encumbers him with help. Analogies in the support of modern science suggest themselves.

Not the least of the author's problems is the dedication. A spouse, mistress or auntie can at least be relied on to accept. (I have even seen, indeed contributed to, a multi-author volume in which the editors dedicated their collection of other men's flowers to their wife.) The most embarrassing of undesired and undesirable favours is the *Festschrift*. Some years ago an aging and eminent physicist, sensible of the approaching danger, declared in an article with the title of "Fest me no Schriften" that he did not wish to become a pretext for the unloading of unrefereed and possibly otherwise unpublishable contributions by his former students and colleagues (whom he probably did not like anyway). The Duke of Wellington once declined a dedication, explaining that he had been obliged to make a rule to do so "because, in his situation as Chancellor of the University of Oxford, he had been much exposed to authors".

As to the foreword, its object is to praise the book, and a hypocritical show of modesty can be avoided if the services of an elder of the subject can be secured. The leaden phrases that fall so easily from the pens of such illuminati are all too familiar. An ingenious ploy was contrived by Newton, who when President of the Royal Society set up a hand-picked committee to prepare a report on the rival claims of Leibniz and himself to the calculus, and then himself wrote an anonymous and fulsome foreword, commending the conclusions. But let us end with a more wholesome example, that of Boswell, who had no need of specious humility in launching his *Life of Johnson*. In what he candidly termed the "Advertisement" to the first edition, he begins: "I at last deliver to the world a work which I have long promised" and concludes, "nor will I suppress my satisfaction in the consciousness, that by recording so considerable a portion of the wisdom and wit of 'the brightest ornament of the eighteenth century', I have largely provided for the instruction and entertainment of mankind". But then his subject was better than the oxides and oxyacids of sulphur.

A pilgrim's progress

W.F. Bynum

The Correspondence of Charles Darwin, Vol.1 1821-1836.

Edited by Frederick Burkhardt and Sydney Smith.

Cambridge University Press: 1985. Pp.702. £30, \$37.50.

A Calendar of the Correspondence of Charles Darwin, 1821-1882.

Edited by Frederick Burkhardt and Sydney Smith.

Garland: 1985. Pp.690. \$100, £80.

Darwin Pedigrees. By R.B. Freeman.

R.B. Freeman, University College London: 1984. Pp.84. £8.

IN THESE contracting times, there is comfort in the appearance of the first volume of a set which, the publishers tell us, will take a score of volumes and most of the rest of the century to complete. The project to locate, edit and publish the complete correspondence of Charles Darwin has already been underway for a decade, supported by large grants from several American funding bodies, including the National Science Foundation, the National Endowment for the Humanities and the Andrew W. Mellon Foundation, and overseen by both British and American advisory committees. The editorial team has already involved more than 20 scholars, designated by a hierarchical variety of titles: editors, associate editors, assistant editors, research associates, research assistants and managing editor. The team has a decided Anglo-American flavour, neatly encapsulated by the two senior editors: Frederick Burkhardt, the American scholar who has masterminded the whole project, and Sydney Smith, the Cambridge zoology don who has done so much over the years to secure the magnificent Darwin archives for the Cambridge University Library. The Darwin family has always been exceptionally cooperative and generous in their dealings with Charles Darwin's Alma Mater, and it is pleasing to see the family represented on the British advisory committee through Darwin's grand-daughter, Lady Nora Barlow.

The Correspondence of Charles Darwin, then, has been, and will be, expensive to produce. Will it be worth it? After all,

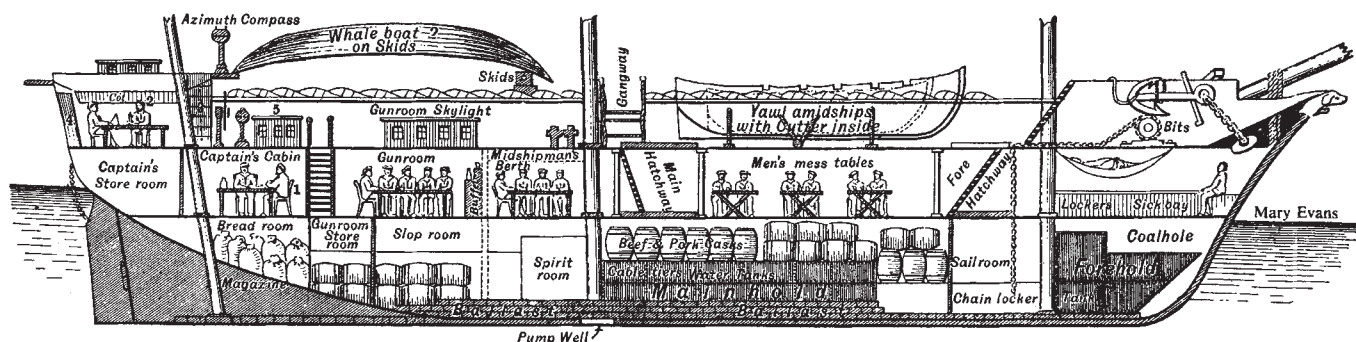
Francis Darwin included a number of the more important letters in his three-volume biography of his father (1887), which he followed up by *More Letters of Charles Darwin* (two volumes, 1903). Charles Darwin's daughter Henrietta Litchfield edited additional ones in her *Emma Darwin: A Century of Family Letters* (two volumes, 1905). Nora Barlow has given us the Darwin-Henslow correspondence (1967) and other scholars have published additional bits and pieces. With about 2,800 letters to and from Darwin already in print, is this present project worth the time and money? The answer is short and unqualified: yes. In the first place, 2,800 letters represent but a fraction of the 13,925 the group has already identified and calendared, and many more will undoubtedly surface.

A Calendar of the Correspondence of Charles Darwin, 1821-1882 gives us a foretaste of the pleasures in store during the coming years and will immediately prove a useful reference tool to historians working on many aspects of nineteenth-century science. The letters are listed and described, and their contents briefly summarized. They come from more than 200 institutions and private collections, and a biographical register lists over 4,000 correspondents and individuals mentioned in the series. Because Darwin lived for so many years in the relative isolation of Down House, he was probably more dependent on the postal service than any other great scientist. The output is staggering: more than 500 letters survive for 1862, almost 600 for 1874 and even

in 1881, with his health visibly failing, the surviving correspondence tops 600 items. The last letter listed in the *Calendar* was written two days before he died.

Darwin, however, was not simply an ardent cultivator of the epistolary art, he was also a distinguished one. His letters are often filled with memorable phrases and shrewd insights and are a joy to read. In no future volume of the complete correspondence are the delights likely to be greater than those contained in the first. This is partly because the time span is longer, from the first surviving letter, written to the 12-year-old Darwin by Mary Congreve, to a letter of 30 December 1836 from Robert FitzRoy to Darwin, discussing publication plans for the diaries, journals and reports prepared in conjunction with the voyage of the *Beagle*. Sandwiched between are more than 300 letters to and from Darwin, and a sprinkling of letters about or affecting him (such as George Peacock's letter of August 1831 to J.S. Henslow, asking him to recommend a naturalist to accompany FitzRoy). They detail a pilgrim's progress, from a spirited young lad variously called Mr Charles, Bobby and Charley, to a mature naturalist, settled in London to work through the large collections of specimens (zoological, botanical, mineralogical and palaeontological) he had acquired on his five-year voyage around the world. The Darwin that emerges is not radically different from the one which numerous biographers and historians have described, but the assembled correspondence does present an uncluttered vision of his early personality, development, and family and social relationships. Two or three points are worth stressing particularly.

First, from the earliest years of the correspondence, Darwin was investing great quantities of libidinal energy engaging with Nature: partly in hunting and fishing, but also in collecting and studying the objects of the natural world. A long, important and largely previously unpublished series of letters between Darwin and his cousin William Darwin Fox, makes clear that insects were Darwin's first scientific passion: "I am dying by inches, from not having any body to



1. Mr. Darwin's Seat in Captain's Cabin

3. Mr. Darwin's Chest of Drawers

4. Bookcase

2. Mr. Darwin's Seat in Poop Cabin with Cot slung behind him

5. Captain's Skylight

Side elevation of HMS Beagle, 1832. Doubts about her size (90 feet in length, 242 tons) were a recurring feature of Darwin's letters prior to the voyage; thus (to W.D. Fox): "My objection to the vessel is its smallness. . . . As to its safety I hope the Admiralty are the best judges".

talk to about my insects", he wrote in 1828; in 1831, on hearing he was off on the *Beagle*, another collecting friend Frederick Watkins wrote to Darwin, "Woe unto ye Beetles of South America, woe unto all tropical butterflies". Despite Henslow's inspiring Cambridge lectures in botany, and despite Darwin's brief geological tour of Wales with Adam Sedgwick, the Darwin who set out on the *Beagle* was first and foremost an entomologist. He very quickly became much more of a geologist, partly through the influence of Lyell's *Principles of Geology* (three volumes 1830-1833), the first volume of which FitzRoy gave him just

before the *Beagle* set sail (the other volumes being sent to him in South America), but mainly through the sheer impressiveness of the geological formations and activity the voyage offered. "There is nothing like geology", he wrote home in 1834; "the pleasure of the first days partridge shooting or the first days hunting cannot be compared to finding a fine group of fossil bones, which tell their story of former times with almost a living tongue."

Darwin's transformation from beetle collector to rounded naturalist is strikingly demonstrated in the volume. Equally significant, however, is the light it sheds on his circle: family, more distant relations and friends. The letters remind us that the sage of Down House was once a carefree young man who lay in the strawberry patches and played Housemaid and Postilion with Fanny Owen. His letters to her are lost, but Fanny's to Charles survive (mostly with strict instructions for him to burn them instantly once he has read them). One wonders what he felt when, barely a year at sea, he received his sister's letter informing him that the girl he described as "la belle Fanny", "the prettiest, plumpest, Charming personage that Shropshire possesses" was engaged to another. Fanny's letters suggest Charles was infinitely better off with Emma Wedgwood, still a shadowy figure in Volume 1. Her time will undoubtedly come in Volume 2.

This lapse into the language of romance is entirely appropriate, for these letters reflect a real world very like the one created by Jane Austen. Darwin's sisters took it in turn to write him each month, and their long and newsy letters are full of the stuff of the social history of upper middle class, provincial England during the 1830s. They send him news of the Reform Bill and the cholera epidemic, but also of engagements and marriages, births and deaths, balls and visits. His sisters despair as the voyage stretches from two to three to five years; and Darwin's own stiff upper lip begins to quiver as seasickness persists, the novelty fades and (after the visit to the Galapagos Islands), the scientific returns decrease exponentially. "In short, I am convinced



The Wedgwood connection — Susannah (1765-1817), daughter of Josiah Wedgwood the founder of the pottery dynasty, and Charles Darwin's mother, with her brother (centre).

it is a most ridiculous thing to go round the world, when by staying quietly, the world will go round with you." He penned this prophetic sentiment in the penultimate letter he wrote home, in July 1836.

This volume is teeming with life and exuberance. There is little of the ill health that was to plague Darwin's later years, and, because his correspondents are allowed to speak for themselves, much about his early circle, from his irascible and fascinating brother Erasmus to his

quietly supportive father to his wonderfully demonstrative sisters. The letters are finely edited and presented, even if the explanatory editorial apparatus could profitably have been fuller. Several appendices provide additional information for these years, including the main events of his life, a list of books and people on board the *Beagle*, and biographical details of those with whom he corresponded or who are mentioned in the letters. The family connections can now be additionally sorted out with the aid of Richard Freeman's new book, *Darwin Pedigrees*. This reprints the scarce pedigree first produced in 1888 by H.F. Burke,

and incorporates Freeman's own researches, particularly on later Darwins and on female members of the family.

For some years now, historians have referred (sometimes disparagingly) to the Darwin Industry. All it needed, it seems, was an injection of foreign capital to show itself alive, well and an immensely fruitful activity. □

W.F. Bynum is at the Wellcome Institute for the History of Medicine, London.

Happy repitition

David L. Hull

The Survival of Charles Darwin: A Biography of a Man and an Idea.

By Ronald W. Clark.

Random House/Weidenfeld & Nicolson: 1985. Pp.449. \$19.95, £14.95.

BIOGRAPHIES of Charles Darwin have become as stylized and predictable as kabuki theatre — Darwin's comfortable childhood, his lazy days at Cambridge, the happy accident that put him on board the *Beagle*, his return to England after his arduous voyage, his marriage and retreat to Victorian gentility at Down House, his mysterious illness, the long gestation of the *Origin of Species*, Wallace's bolt from the blue, T.H. Huxley taking on "Soapy Sam" at the 1860 meeting of the British Association for the Advancement of Science, the eventual triumph of Darwin's theory, Darwin's later publications, his death and burial at Westminster Abbey. Clark's biography is no exception, although he does strew this familiar narrative with interesting titbits. For example, anyone reading Darwin's *Life and Letters* might notice that Darwin called his residence near the village of Downe in Kent "Down House", instead of "Downe House", and wonder why. Clark explains that the village was "Down" until 1842 when the government added the "e" to distinguish the village in England from the

county in Ireland. There was, after all, the continuing Irish Problem. Darwin refused to adopt the new spelling for his own house.

Clark does not stop with Darwin's death in 1882. His book is the biography of a man, but it is also the biography of an idea. He traces the decline of evolutionary theory towards the end of the nineteenth century, its rough treatment at the hands of such Mendelians as William Bateson, the resurgence of Lamarckism, the "Monkey Trial" in Tennessee, the increasing influence of mathematics on the theory and the formation of the Modern Synthesis. He concludes with two chapters, one entitled "Molecular Biology Takes Over" and a second in which he describes recent controversies over the tempo of evolution and over cladistic analysis. These discussions are necessarily somewhat cursory, and some evolutionary biologists might protest that molecular biology has not taken over their field. The most that can be said is that molecular data have been incorporated into traditional views of the evolutionary process. With respect to the pace of evolution, Clark shows that emphasis on slow, gradual evolution has alternated periodically with more saltative views. Eldredge and Gould's punctuational model is but one more swing of the pendulum.

Another theme that recurs in the book is the social responsibility of scientists. Scientists do not publish in a social vacuum. Other experts read their work but so do theologians, reporters, high-school teachers and politicians. For example, the

anonymous reviewer of Darwin's *Descent of Man* (1872) in the London *Times* chastized Darwin for being "reckless" in publishing such "disintegrating speculations" in such perilous times. Fifty years later and an ocean away, Bateson was to be given a similar lesson when he claimed at a meeting of the American Association for the Advancement of Science that the "origin and nature of *species* remains utterly mysterious". As might be expected, newspaper headlines declared triumphantly "Darwinism Disproved".

Another 50 years on the scenario was replayed when critics of sociobiology warned that statements made by E.O. Wilson and company could easily be used by racists and

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Down House, at Downe in Kent, Darwin's home from 1842 until his death in 1882.

sexists. These criticisms in turn were taken up by Creationists as more proof that evolutionary theory is false. Even more recently, advocates of a new theory of biological classification termed "cladistics" have proclaimed that cladistic classifications are not about evolution but about patterns of character distributions, no matter how they may have arisen. Rumours of Darwin's death in South Kensington were a result. More recently still, Tom Bethell in *Harper's* has complimented cladists for having a truly scientific outlook. "Only evolutionary agnostics like Patterson and Nelson and other cladists seem willing to live with doubt. And that, surely, is the only truly scientific outlook." As far as Bethell is concerned, the proper attitude towards evolution is the one that Huxley advocated for the existence of God — not proven. On these standards, very little has been proven in science, including relativity theory, quantum theory, not to mention continental drift. The authenticity of the Shroud of Turin is quite another matter. There is no doubting here.

Once again, Creationists are sure to make ample capital out of this intramural dispute among scientists. But what are scientists to do? No matter how judiciously they present their views, no matter how technical the journal, their statements can always be misused by those whose political or religious convictions outweigh considerations of intellectual integrity. If nothing else Clark's history of Darwinism shows that scientific disputes continue, as do marriages and divorces according to Dr Pangloss, in happy repetition. □

David L. Hull is a Professor of Philosophy at Northwestern University, Evanston, Illinois.

The conversation continued

Philip Kitcher

The Nature of Selection: Evolutionary Theory in Philosophical Focus.

By Elliott Sober.

MIT Press:1984. Pp.383. \$33.75, £30.

VIRTUALLY no philosopher is immune to the fear of irrelevance. Despite the reassurance that one is grappling with problems that have challenged the Western intellectual tradition since the time of the Greeks, there is always the quiet voice that asks "Does anyone care?". So it is tempting to announce that philosophy is — and perhaps has always been — the servant of the sciences, and to immerse oneself in the theoretical problems that arise within some more self-confident domain of inquiry. Even this may not put an end to the whispered doubts. If philosophers turn meta-scientists, do they manage to contribute anything that a reflective practitioner could not have achieved more easily? Is philosophy only useful when it tidies the corners that talented scientists are too busy to set in order for themselves?

A long and distinguished tradition in the philosophy of science reveals the impact of combining philosophical understanding with deep knowledge of the sciences. The tradition is strongest in the philosophy of physics, where reflections on space, time, cause and chance have proven valuable both to science and to philosophy. More recently, there has been a fruitful alliance between philosophy and psychology, in which students of cognitive science have looked to philosophers for conceptual guidance, offering in return new perspectives on old problems about mind, language and knowledge. The publication of Elliott Sober's new book shows that biology's turn has now come.

On reading a draft of *The Nature of Selection* in the autumn of 1983, I formed the opinion that it would illuminate both biology and philosophy. That impression is amply confirmed by study of the published version. Sober has focused evolutionary theory for philosophers, and thereby provided new material for inquiry into the general problems of the philosophy of science, the problems of explanation, confirmation and theory structure. He has also addressed a number of major conceptual problems that currently confront practising biologists. In his hands, some disputes about natural selection are reformulated; others are dissolved.

Sober's main project is to understand the causal notions that figure in evolutionary theory. He starts from the conviction that many contemporary debates about the evolutionary process persist because the protagonists have failed to understand the character of the causal claims that they and

their opponents are making. The first part of the book is devoted to an analysis of the concepts of selection, adaptation and fitness, and to a delineation of the role that these concepts — as well as the concept of chance — play in evolutionary theory. The second part develops the analysis in the context of a dispute to which Sober has made substantial contributions, the controversy about units of selection. Substantial though that previous work was, the present discussion supersedes it.

On Sober's account, evolutionary theory is best construed as a theory of forces. Evolution is a process in which typically (though not invariably) gene frequencies change along a lineage, and evolutionary theory attempts to specify the forces that produce the changes. An immediate consequence of this perspective is that the old problem about the evolutionary tautology is broached and quickly disposed of. (Those who are tempted to use the term "tautology" in discussions of evolution ought to be required to read Sober's second chapter.) A more important result is the distinction between selection *of* and selection *for*. Selection of *objects* contrasts with selection for *properties*. The distinction can be made by considering the electorate of the United States. American voters appear to select candidates who are bereft of foresight and understanding. Yet there is no selection for lack of foresight. The property selected for is surely telegenicity.

Although there are several other discussions in the first part of the book that will interest scientific readers, it is probable that they will think of the second part as where the action is. Sober's approach to the problem of units of selection starts with a chapter, allegedly designed to alienate all parties to the controversy, in which he tries to uncover confusions and fallacies in all the major positions. He continues with an analysis that makes explicit appeal to the concept of cause: a necessary condition for X-selection (group selection, genic selection, species selection and so forth) is that there should be some property of Xs that is a positive causal factor on the survival and reproduction of organisms. (Sober eventually generalizes, allowing for entities other than organisms to serve as "benchmarks" of selection.)

However, the concept of cause cannot be

● The Royal Historical Society has published a collection of correspondence to complement Jack Morrell and Arnold Thackray's *Gentlemen of Science: Early Years of the British Association for the Advancement of Science* (published in 1981 by Oxford University Press and reviewed in *Nature* 294, 19; 1981).

In *Gentlemen of Science: Early Correspondence* Morrell and Thackray have gathered together "what [they] judge to be the 294 letters of greatest significance to the student of early Victorian science and culture", focusing upon letters written by and to William Vernon Harcourt, a pivotal figure in scientific matters during the early and middle nineteenth century.

taken for granted, and Sober invokes the philosophical literature on probabilistic causation to provide a more detailed analysis. His final suggestion is that *C* is a positive causal factor in the production of *E* in a population, if, for things in the population, the occurrence of *C* does not lower the probability of occurrence of *E* on any background of causally relevant conditions, and raises it on the background of at least one causally relevant background condition. (Roughly: smoking is a positive causal factor in the production of heart attacks if, whatever constellation of traits an individual has, smoking would make that person more prone to have a heart attack.) Sober is aware of the limitations of this analysis, which he forthrightly describes as "circular" (p.291). It would be more appropriate to treat the proposal as a recursion clause for extending our ability to identify positive causal factors. For it is this extended ability that we need to tackle the controversies about units of selection. I shall postpone the question of whether the "circularity" is bothersome.

The last chapter brings the philosophical account of causation to bear on four debates about levels of selection. Sober offers analysis of Dawkins's celebrated rowing analogy, and a diagnosis of why we should not be rushed into thinking that the gene is the one and only unit of selection. However, Sober is careful to point out that his arguments help us understand the idea of "selfish DNA" (developed by Doolittle, Sapienza and others). He also attempts to provide conceptual underpinnings for the claims about macroevolution and "species selection" that have been advanced by Stanley, Gould, Eldredge and Vrba. Here, too, is a discussion of Hamilton's ideas about inclusive fitness and local mate competition, and an extended exploration of proposals about group selection.

The account of inclusive fitness seems to me to be the least useful part of this chapter. Sober starts by suggesting that Hamilton "consolidated an insight that had been available within population genetics for some time" (p.336). Quickly adding that he does not mean that Hamilton underwrote Haldane's apocryphal quips about sacrificing himself for two brothers or eight cousins, Sober continues by reminding us that the fitness of a trait is standardly defined as the average fitness of the bearers. So Hamilton's insight can be reduced to a rather elementary idea: "There are many ways for a gene to be advantageous *on average*" (p.336).

This evaluation strikes me as parallel to the view that Darwin consolidated an insight of Malthus. Hamilton does not seem primarily interested in identifying the fitnesses of individual alleles or of allelic pairs. (That approach was later adopted by Dawkins, and is the target of much of Sober's criticism.) Inclusive fitness was introduced because of the apparent inadequacy of the classical notion of fitness for

phenotypic traits, an inadequacy that is most obvious in traits that reduce the expected reproductive success of their bearers. When we average over the set of individuals with any such phenotype we are likely to find that the (classical) fitness lies below the mean for the population. For the lucky recipients, the kin whose reproductive success is boosted by the fact that others have the phenotypic trait in question, *are not themselves bearers of the trait*. They only enter into the set over which it is appropriate to average when we compute the chances for spreading the gene(s) that underlie the phenotypic trait, and thus calculate — as Haldane did in rough-and-ready fashion, and as Hamilton did systematically — the way in which reduced genetic representation through descendants is offset by increased genetic representation through kin.

In general, I find Sober's discussion of inclusive fitness confusing because his use of "altruist" can cover both individuals who bear genes that might (in the right circumstances or in the right combination) dispose them to sacrifice their own reproductive interests to those of others, and individuals who actually do make these sacrifices. The treatment of group selection is another story. Sober's analysis of this tricky notion is one of the high points of the book.

The distinction between group and individual selection is introduced with an apt example. We are to imagine a set of internally homogeneous populations with different average heights. We observe that tall individuals are reproductively more successful than short ones. We note also that groups of taller individuals leave more descendants than groups of shorter individuals. How do we tell if we have a case of individual selection or a case of group selection?

Everything turns on whether the causally relevant factor is the property of being tall or of being in a group with a large average height. To recognize the difference, we shall need to consider a larger ensemble of populations, asking how short individuals would have fared if they had found themselves in tall groups, and so forth. The probabilistic treatment of causality allows the conditions to be stated precisely.

Sober's work will have been done if the controversies about group selection can now be resolved by appeal to the empirical evidence. There is one obvious possible snag. The probabilistic conditions bear reference to causally relevant background conditions. If reductionists and anti-reductionists disagree on the properties that can be invoked as genuine causally relevant background conditions, then Sober's analytic apparatus may fail to bring resolution. Faced with a putative case of group selection, the reductionists produce what they are happy to count as a causally relevant property of individuals, given which the group characteristic makes no difference to the probability of survival

and reproduction. If their opponents contend that the alleged property is contrived, an artefact rather than a genuine cause, then Sober's analysis will have to be extended by explicating the notion of a causally relevant background condition. So we cannot tell in advance if all the conceptual work has been done. But even pessimists should be confident that Sober has taken us a long way forward.

The Nature of Selection is not merely a book that clarifies dramatically some vexed issues. It also points to new questions and, perhaps, to refinements of causal notions that it introduces. In its achievement and in its open-endedness it deserves comparison with the work that Sober singles out as having started his own train of thought. George Williams's *Adaptation and Natural Selection* has been important to philosophers as well as to biologists for its combination of biological insight and philosophical acumen. Sober's is the answering philosophical voice, the voice of a first-rate philosopher and a knowledgeable student of contemporary evolutionary theory. His book, like Williams's, merits broad attention among both communities. It should also inspire others to continue the conversation. □

Philip Kitcher is Professor of Philosophy at the University of Minnesota. His latest book, Vaulting Ambition: Sociobiology and the Quest for Human Nature, will be published in September.

Advertisement

ANNOUNCING NEW PUBLICATIONS

Continuing the Nuclear Dialogue Selected Essays

by

ALVIN M. WEINBERG

and

Nuclear Criticality Safety

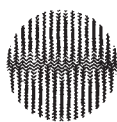
by

RONALD A. KNIEF

order today
from the

American Nuclear Society
555 N. Kensington Ave.
La Grange Park, IL 60525
(312) 352-6611

Reader Service No.46



INRA

PUBLICATIONS

Just published

Animal Production

Le Cheval : reproduction, alimentation, exploitation

1984, 692 p. 250,00 F
Prix de lancement

Physiologie et pathologie périnatales chez les animaux de ferme

1984, 474 p. 195,00 F

Insémination artificielle et amélioration génétique : bilan et perspectives critiques

Colloque Toulouse, novembre 1983
1985, 368 p. 140,00 F

Les maladies de la chèvre. Goat diseases.

Colloque international, Niort, octobre 1984
1984, 750 p. 270,00 F

The reproductive potential of cattle and sheep. Le potentiel de reproduction des bovins et des ovins.

Joint Israeli-French symposium, Rehovot (Israël), February 1984.
1984, 413 p. 150,00 F

Plant Production

Les essais de fertilisation de longue durée de la station agronomique de Grignon - Résultats expérimentaux (période 1938-1982)

R. MOREL, Th. LASNIER, S. BOURGEOIS
1984, 335 p. 120,00 F

Physiologie du maïs. Colloque Royan
A. GALLAIS éd.
1984, 574 p. 240,00 F

La conservation des céréales en France. Compte-rendu d'actions pluridisciplinaires, 1980-1982

1984, 171 p. 60,00 F

Les cultures hors-sol

1985, 410 p. 130,00 F

Petite paysannerie en France

G. JEGOUZO
1984, 232 p. 145,00 F

Variation et variabilité des agents phytopathogènes. 26^{ème} colloque de la Société française de Phytopathologie.

Avignon, mai 1984.
1984, 235 p. 90,00 F

6^e conférence européenne et méditerranéenne sur les rouilles des céréales. Grignon, septembre 1984.

1984, 223 p. 110,00 F

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

Service des Publications - Rte de St-Cyr

78000 Versailles

Télex : INRAPUB 699368F

Larger than life

R.C. Lewontin

On Being the Right Size and Other Essays.

By J.B.S. Haldane. Edited by John Maynard Smith.

Oxford University Press: 1985. Pp. 187.
Pbk £4.95, \$6.95.

I MUST confess that when I first met J.B.S. Haldane I thought him a bit of a humbug. Dressed in his *dhoti*, smoking a twisted black Burmese cigar and addressing a group of admirers in a Dutch café, he seemed unreal. Nor was I alone in that opinion. Once in an American department store where we had gone to buy some things for Helen Spurway, Haldane pulled out an Indian rupee note, announcing to the sales clerk "I'm a Hindu, you know". As he left, she turned to me and said "That's funny. I could have sworn he was one of those high-class Englishmen".

But, of course, Haldane was vastly more than a showman. He was the most incisive and wide-ranging intellect I have ever known, original, iconoclastic, a master at making connections between apparently unrelated phenomena, and by far the best popular science writer who has ever addressed an English-speaking readership. All these aspects are displayed in this collection of his essays, chosen with great judgement by his former student and colleague John Maynard Smith, essays ranging from the origin of life and the gods, to the Last Judgement, somewhere around the year 36,000,000. In between Haldane explains in plain English what "hot" means, how the organism is produced by interaction of its genes and its developmental environment, why one can

survive for hours shut up in an air-tight box and how capitalism has made a commodity out of health. Written as clearly as they are, these essays tempt the reader to take them all at one sitting, but I advise against it. The best prescription is one a day followed by deep thought.

Haldane the humbug shows up occasionally, as when he quotes extensively from the Norse epic *Volospa* in the original, a bit of flummery that can only be meant to give an impression of fluency in an esoteric language. Yet it seems so unnecessary. Haldane was a person of genuinely wide learning, as shown, for example, in the essay "Is History a Fraud", where a variety of historical detail is displayed in the service of a central idea: the superiority of the practical over the merely intellectual. It is a sign of the contradiction that marked Haldane's world view that he could prefer Pasteur above Darwin because the former solved a "practical problem of living", yet at another time could praise an idea as "interesting, even if not true".

Haldane the critical scientist comes out best in "The Biology of Inequality" which shows a subtlety of understanding of the relationship between gene, environment and organism that is, even today, absent from the thinking of most biologists. Yet here again is a contradiction. Haldane makes clear in this essay that it is not the phenotype, but the pattern of reaction to a contingent environment that is inherited. Yet earlier, in "The Future of Biology", he writes of "the main gene that determined his father's mathematical powers". Nor was it simply wisdom acquired in the ten years between these essays. To the end of his life, Haldane remained committed to a hereditarianism in regard to human psychic development, a biological determinism that may seem strange in

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Public faces — Haldane addressing a "United Front" meeting in Trafalgar Square, 1937 (right), and at the Indian Statistical Institute, Calcutta. The pictures are reproduced from J.B.S., Ronald Clark's biography which was re-issued in paperback last year by Oxford University Press.

someone who claimed to be a revolutionary socialist. The key to this apparent paradox can be found in what appears to be Haldane's preoccupation with religion.

In essay after essay in this collection Haldane returns to the subject of religion. It is not simply that the editor has chosen essays on religious subjects, such as "When I am Dead", "God-makers" and "Science and Theology as Art Forms". Haldane brings in Christianity, Hinduism, religious myth in general and an Egyptian fakir in particular as integral to the points he is making about history and science. To understand why this is so, we must realize that Haldane was born in the last years of Victoria's reign and his intellect was formed in the Edwardian period. The primary intellectual struggle in his view, a view shared by Marxists of the 1930s, was the struggle between materialism and

philosophical idealism. Marxism for Haldane, as for H.J. Muller, meant the fight to overturn the superstition, both scientific and social, that marked nineteenth century ideology, and its replacement by a materialist, scientific society. Part of that materialism was a belief that biological technology including eugenic measures could be instrumental in putting society on a better basis. The view seems naive to us now. If biology is to be relevant to reorganizing society, it must be the biology of what is in our heads, not what is in our gonads. Nevertheless, we must side with Haldane's desire to replace mystifying social myths by rational realities. □

R. C. Lewontin is a Professor of Biology in the Museum of Comparative Zoology, Harvard University.

Schools of thought

Struther Arnott

Graduate Research: A Guide for Students in the Sciences.

By Robert V. Smith.

ISI Press, 3501 Market Street, Philadelphia: 1984. Pp.182. Hbk \$21.95 (North America), \$24.95 (elsewhere); pbk \$14.95 (North America), \$17.95 (elsewhere).

HERE is a programme instructing the ambitious mediocracy how to apprentice themselves to the trade of being a (US) scientific worker. It begins at the same inspirational level as it continues with the thought that "it is crucial for graduate students in the sciences to master this skill [in research] because their advanced degrees depend upon it". It concludes with the optimistic and self-congratulatory assurance that "you should expect to find a good position".

This is a "how to" manual with a vengeance: how to get started; how to choose a research field and a mentor; how to manage time and libraries; how to write dissertations and papers; how to get grant support and a job; even how to note rather than to take cognizance of and how to prepare a consent form for research on human subjects. Thank goodness the chapters on ethics and creativity are not "how to" ones. But there we are regaled with a re-run of Kekulé's reverie which produced the structure of benzene. Inexplicably absent is the crucial part of the dream where some of the dancing chains of atoms joined ends to form rings. Such persistence in missing the point is a general characteristic of this book.

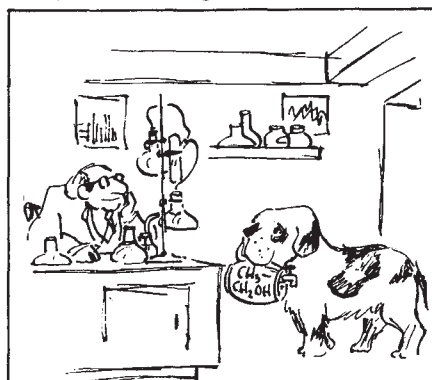
There is a tidy-minded part of me which was impressed by Dr Smith's thoroughness in mustering many useful thoughts about our graduate schools. I and my colleagues surely would want our doctoral students,

potential and actual, to be aware of the management structures in our (usually very large) universities, to be familiar with the various opportunities for financial support, and to be chary of over-involvement in teaching and in passive attendance at courses and seminars. However, not all the wise saws and modern instances are as pleasing. Few of us would dispute that "excessive television is a deterrent to purposeful observation" but would any of us want to be saddled with protégés needing such banal advice or with disciples who would formulate the hypothesis that "sensory deprivation in human subjects is associated with increased alienation and manifested as loneliness" only after a survey comparing deaf and blind widows with widows having normal hearing and sight?

The author does assert at one point that "research... requires certain personal traits and practices. Some characteristics must already be part of the individual, some can be improved with practice". I agree completely. What strains my credulity is the implication that a beginning graduate student typically could have postponed until reading this book the acquisition and development of the traits of a successful scholar or researcher. Yet expectation of perpetual opportunity for educational rehabilitation has been a pervasive ingredient of public policy in the United States for at least a generation. Only recently has there been general recognition of the consequences, such as high schools staffed with inadequately rewarded — and therefore indifferent or demoralized teachers — graduating, not very discriminatingly, 75 per cent of their entering classes. Better paid college professors can thereafter, much more expensively, try to remedy ingrained habits of intellectual indolence. Professors condemned to ushers do not resist this marketplace philosophy nor strive to maintain system in national educational arrangements. The result is loss of rigour, in secondary and in higher education, and buffeting of both by fads and fashions. My own in-

stitution recently suffered from a neap tide of interest in computer science which briefly was the proclaimed major subject of more than half the freshman class in our School of Science. Today, that tide is receding, exposing a floundering population dismayed by its lack of quantitative training, analytical skills and capacity for sustained hard work.

A committee of the American Association of Colleges now has demanded baccalaureate programmes which would ensure that their alumni would reason well, would read, write and speak with distinction,



Reproduced from Science Goes to the Dogs, a new collection of cartoons by Sidney Harris published by ISI Press. Price of the book is \$7.95.

would understand numerical data and the scientific method, and would have historical consciousness and knowledge of the language of the fine arts. Graduate schools, anxious for distinction, will demand at least this, as well as evidence of ability to study in depth which, as the AAC report describes, requires sequential learning, building on blocks of knowledge that lead to more sophisticated understanding and encourage leaps of the imagination and efforts at synthesis.

Now indeed may be the time for us to attend to the quality and configuration of our graduate schools. The main places of employment for holders of the PhD degree are colleges and universities. Population trends will ensure a 15 per cent fall in the traditional university student population by 1995. Scepticism about the value of some academic qualifications could accelerate this decline substantially. A shift in public policy towards investment in earlier stages of education could make the descent precipitous. Even without a revolution in public policy, there may be less than 100,000 academic positions to be filled in the next decade. Success will fall to those graduate schools which order their requirements and incentives so as to be able to recruit more discriminatingly. In such an environment, Dr Smith's book would be useful only to foreigners curious about some of the minutiae of the academic research environment in the United States.

Struther Arnott is Vice President for Research and Dean of the Graduate School at Purdue University, Indiana.

Angles on knowledge

David Miller

The Limits of Science.

By Nicholas Rescher.

University of California Press: 1985.

Pp.225. \$34.50, £31.50.

THE question of whether there are any limits to empirical science is one that, since Kant anyway, has never been far from the centre of the theory of human knowledge. Logical positivists persuaded themselves (and others) that meaningful investigation is restricted to the results of observation, experiment and induction. Less extreme sufferers from scientism have held that there is no knowledge but scientific knowledge. On the other side it has been suggested that the traditionally conceived method of science as one of "rational triangulation from the empirical data" (as Rescher describes it on p.198) sets more of an external limit to science than an internal one.

In this work, much of which is collated from his innumerable previous books, Rescher approaches the issue from a number of angles: whether there are topics wholly beyond the mandate of science ("no" in Chapter 7, "yes" in Chapter 12); whether science could be perfected, even in principle ("no" in Chapter 9); whether there are problems within its domain that science is in principle unable to answer ("no" in Chapters 4 and 8); whether there are significant questions that for practical (technological) reasons science cannot handle ("yes" in Chapter 10); whether science progresses steadily towards theoretical perfection ("no" in Chapter 5); whether it progresses at an instrumental level ("yes" in Chapter 6); whether extra-terrestrial science might be scientifically more advanced than ours (question dismissed in Chapter 11).

The discussion throughout is thoroughly unsatisfactory: superficial, confused, philosophically unsophisticated, repetitive, neologistic, above all deeply antipathetic to theoretical speculation. For example, in the account in Chapter 11 of how our science and alien science might be related, a discussion whose only real conclusion is that there is nothing of interest to say, we read:

A tiny creature living its brief life span within a maple leaf could never recognize that such leaves are deciduous. ... Science is limited to the confines of discernibility: as Kant maintained, the limits of our experience set limits to our science [p.197].

So much for the power of ideas. Yet in Chapter 8 Rescher seems properly to appreciate that science is not confined to what we can experience but a method of going

wildly beyond it. In a similarly glum mood Rescher seems to suppose that because other planets might be physically very different from ours, scientists there might use mathematics... very unlike ours.... their 'geometry' could be something rather strange, largely topological, say, and geared to flexible structures rather than fixed sizes or shapes [p.177].

Just like some of our mathematics by the sound of it.

Can current science tell us anything about future science, about extraterrestrial science, about possible limits to scientific inquiry? The answer seems obviously that it might be able to, but that what it tells us might be wrong. Rescher persistently muddies this issue, gliding from the platitude that "no adequate justification can be found for the view that science has barriers" (p.130) to the assertion that it has no barriers. He seems not to have appreciated that if "present science cannot speak for future science..., [if] there can be no basis for claims of inherent unanswerability" (p.129), then present science cannot speak for present science either, and there can be no basis for its assertions about the natural world. Yet Rescher is unwavering in his belief that the results of science can be justified. We are told that the advance of science amounts to obtaining "a firmer warrant for our claims", not "more of the truth"; and we are warned of the "fallacy... that moves from a picture-of-

nature's being a *better-warranted picture* to its being a *better picture*" (pp.73-74). Only someone who confuses justification with truth, and thinks that all moves have to be justified, could see this move as a fallacy rather than as an attempt to discover the truth.

The question of the existence of insolubilia is itself eventually resolved in Chapter 8 as follows:

How could we possibly establish that a question *Q* will continue to be *raisable and unanswerable* in every future state of science...? ... [If we] argue that the answer to *Q* lies 'in principle' beyond the reach of science, ... it is difficult to see how we could maintain it to be an authentic scientific question [p.128].

In other words, science has no limits except where it has limits.

Most of the topics treated in this book are treated more than once, as though the different chapters had originated from different authors. The writing is always verbose, and horrible hyphenated words ("issue-killer" on p.17, "question-disallowing" on p.22, "hole-internal" on p.131, for example) abound. On p.52 the same 130-word quotation appears twice: in the text and in a footnote appended to it. The word-processor has a good deal to answer for. In most other respects the book is well produced. □

David Miller is Senior Lecturer in Philosophy at the University of Warwick.

Word and number meta-magic

David E.H. Jones

Metamagical Themas: Questing for the Essence of Mind and Pattern.

By Douglas R. Hofstadter.

Basic Books: 1985. Pp.852. \$24.95. To be published in Britain later this year by Viking.

DOUGLAS Hofstadter is widely and justly celebrated for his book *Gödel, Escher, Bach*, and for his co-production with Daniel Dennett of *The Mind's I*. Between June 1981 and September 1983 he succeeded Martin Gardner in running *Scientific American's* "Mathematical Games" department. He gave his contributions the anagrammatical title of "Metamagical Themas". All 26 of these essays are reprinted in this book. Each now has a postscript of subsequent reflections and developments; in addition there are seven new, specially-written "Themas".

Hofstadter claims that his Themas form a scatter of random dots around his intellectual "home territory", which is concerned largely with thinking, thinking about thinking, and thinking about language. There are Themas on self-referring and self-replicating sentences, and

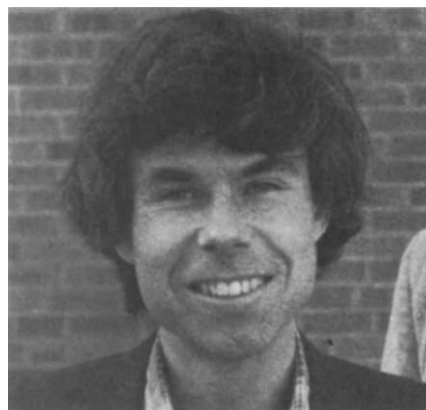
ones on the search for consistent patterns and styles in music and in geometrical art-forms. Hofstadter's own research into artificial-intelligence programming is reflected in several Themas, such as those on the pattern-recognition of sameness underlying difference, in particular the recognizable sameness of letters of the alphabet throughout unimaginably many variations of form and fount.

Other Themas continue the tradition of "Mathematical Games", dealing with such topics as Rubik's cube, the mathematics of chaos, computer languages, psychological games with numbers, and logical dilemmas and paradoxes. A rather special cluster of Themas locates a societal region in Hofstadter's territory: his concern over mass gullibility, innumeracy and some political issues. Each Thema is complete in its own right, but most are buttressed and extended by others elsewhere in the book. It's a format ideal for browsing; rather like his previous books, in fact. There is an excellent bibliography, with short notes on the works cited.

To illustrate Hofstadter's style and matter, I'll outline one Thema in greater detail, No.23 "On the Seeming Paradox of Mechanizing Creativity". Could a computer be creative in the human sense? Or must it always be, in effect, a huge fast-scanned dictionary of clichés? After considering the background to this question, Hofstadter introduces an example of

● P.B. Medawar's *The Limits of Science*, first published in the United States by Harper & Row, and recently in Britain by Oxford University Press, was reviewed in *Nature* 312, 203 (1984).

extreme uncreativity — the blind repetitiveness of instinctive behaviour shown by the wasp *Sphex*. Typically, he gives this behaviour a name: sphexishness. Its opposite, mental flexibility, he calls anti-sphexishness; and to illustrate the notion he positions various machine, animal and human behaviours along a continuum connecting these extremes. The essence of anti-sphexishness, says Hofstadter, is the ability to break out of loops and ruts of all



Hofstadter — love of word-play.

kinds. He has a word for this, too: "jootsing" (jumping out of the system). It depends, he argues, on watching what you're doing, so as to break up any routine which is getting repetitive. This leads into his favourite topic of self-watching and self-referencing systems, recursive theorems and computer-routines, and the use of recursiveness in "creative" computer programs. He ultimately concludes that some sort of Gödel theorem forbids any exhaustive formalization of creative ability. Any conceivable creative genius or artificially intelligent program must ultimately remain a prisoner of its own style — just like the humble *Sphex*.

In the postscript to this Thema, Hofstadter throws out some intriguing remarks about the subtleties of "style" (What makes a tune catchy? Why do trade-names such as "Day-glo" and "Turbo-matic" now seem vaguely dated?), and comments on a new rival artificial-intelligence strategy inspired not by recursion but by the ideas of statistical mechanics. The reader is directed to related discussions in other Themas.

Hofstadter is no mere armchair theorist. He constantly supports his ideas by appeals to practical experience or to intriguing psychological experiments. Expounding the Turing test for computer intelligence (can you distinguish the computer from a human being by its responses to questions?), he presents a transcript of an actual Turing test, with a delightful surprise ending. His discussion of the "Prisoner's Dilemma" paradox, a sort of mathematical microcosm of the tension between competition and cooperation, is enlivened by accounts of several experimental studies. Two of them were tests carried out by Hofstadter himself on human subjects, the other two were tournaments between computer pro-

grams. The tournaments marvellously illuminate the biological concept of the "evolutionarily stable strategy" — the optimum behavioural strategy towards which a group of competing organisms will evolve over the generations. This is used to suggest how socially cooperative creatures could have evolved from a primitive population of warring egotists.

Hofstadter's style of writing has its own pleasures and pleasantries. He loves playing with words; puns, anagrams and literary jokes abound. This almost dangerous infatuation with word-play at times works against him. Three whole chapters are devoted to self-referencing sentences, a self-indulgence quite unjustified by the limited conclusions reached. Another chapter presents numerous examples of literary nonsense, again without any moral that I can see. Set against the commentaries by (for example) Aldous Huxley, George Orwell and John Updike on nonsense-verse considered as a penetrating subversion of the stifling world of convention, this is very thin gruel.

The most serious lapses, though, are political. Hofstadter declares himself a natural activist, easily "fired up" by various good causes. The ones he espouses here, those of nuclear disarmament and feminism, are badly served by his simplistic enthusiasm. His anti-nuclear Thema is a sadly flat and contrived parable, arguing that if enough people devote enough time to denouncing the arms-race, it will go away. Once world leaders realize the complete symmetry of the U.S./S.U. confrontation (there's that word-play again) all will surely be well. His feminist Themas point out that in normal grammatical English usage, "he" can include "she" and so on. He seeks to make this a matter for resentment among women, and guilt in men, by means of an extraordinary attempted mapping from a region of very high guilt-density, that of race relations. Orwell, in his celebrated essay "Notes on Nationalism", remarks that the political bigot's great sensitivity to specific words and phrases probably has overtones of word-magic — the belief that the world can be changed by proper incantation. Despite his fascination with words, Hofstadter is too rational a soul to be a bigot. He admits that he may be over-acting. Certainly all previous attempts to change society by word-manipulation have merely enriched the language's glittering treasure-house of euphemisms and polite evasions.

But this is to carp. Most of Hofstadter's Themas are concerned with his areas of special knowledge and insight, and sparkle with linguistic fun, provocative insights and intellectual challenges. This random walk through his intellectual "home territory" is highly recommended. □

David E.H. Jones is a guest staff member of the Physical Chemistry Department, University of Newcastle upon Tyne, and a science consultant to industry and the media.

Life with a lighter touch

Martin Gardner

"Surely You're Joking, Mr. Feynman!": Adventures of a Curious Character.

By Richard P. Feynman with Ralph Leighton.

W. W. Norton: 1985. Pp.350. \$15.95. To be published in Britain later this year by W. W. Norton.

"THE stories Feynman tells about himself would make a book", an interviewer wrote in 1963. Here, 22 years later, is that book. It is not an autobiography. You will learn nothing from it about the work on quantum mechanics that earned Richard Feynman a Nobel prize in 1965. What you will learn is a great deal about the flamboyant personality of a great theoretical physicist: his amazing range of interests, his unusual hobbies, his enthusiasms and animadversions, his dislike of pompous fools, his unpredictable behaviour and, above all, his fondness for outrageous comedy.

Feynman's famous three-volume *Lectures on Physics*, and his marvellous little book *The Character of Physical Law*, were based on taped lectures. *Surely You're Joking, Mr. Feynman!* is a taping of Feynman's uninhibited conversations over the years with his friend Ralph Leighton. "That one person could have so many wonderfully crazy things happen to him in one lifetime is sometimes hard to believe", Leighton writes in his preface.

The reason they happened is that Feynman has a knack of creating his own adventures. When, for instance, he was a young man at Los Alamos, working on the top secret bomb, he discovered a hole in the perimeter fence. Anyone else would have reported this to the authorities and that would have been the end of it, but Feynman is incapable of passing up a chance for a lark. Out of the main gate he goes, back through the hole, then out again. He keeps this up until the sergeant at the gate suddenly realizes that some strange character is always going out but never coming in.

Feynman's expert ability to pick locks followed naturally from his lifelong passion for puzzles. To demonstrate how loose security was at Los Alamos, he would secretly open the safes of top scientists, leaving little notes that said such things as "I borrowed document LA4312 — Feynman the safecracker". But such idiosyncracies were not appreciated by everyone, least of all the military. When three psychiatrists tried to test Feynman for Army service they quickly found themselves being tested. "How much do you value life?", one of them asked. "Sixty-four", Feynman answered. The Army's final verdict: mentally deficient. (The chapter on this banter, headed

"Uncle Sam Doesn't Want You!", is one of the book's funniest.)

My favourite anecdote, however — it's vintage Feynman — concerns a joke his friends once tried to play on him. In Japan Feynman had learned some Japanese, and in Portugal he gave lectures in Portuguese. At a party someone thought it would be amusing to see how this "man of a thousand tongues" would react if a Caucasian lady, who grew up in China, greeted him in Chinese. "Ai, choong, ngong jia!" she said with a bow. Taken aback, Feynman swiftly decided the best thing to do was imitate the sounds she made. "Ah, ching, jong jien!" he replied, returning the bow. "Oh, my God!" the lady exclaimed. "I knew this would happen. I speak Mandarin and he speaks Cantonese!"

From childhood on, anything mysterious or puzzling instantly aroused Feynman's curiosity. Growing up in Long Island, New York, he taught himself how radios work and he became the neighbourhood's youngest radio repair man. He invented ingenious experiments with ants to work out how their brains were programmed. Intrigued by hypnotism, he allowed himself to be hypnotized. To experience out-of-body hallucinations, he floated in a sensory deprivation tank. Always, wherever he was or whatever he did, the wheels in Feynman's brain never stopped whirring. Every experience posed new challenging questions. What goes on here? Why does this work? Can it be done better? The best example echoes Newton and the apple (no, that story isn't a myth). Feynman once watched someone toss a plate in the air and noticed that its wobble went around faster than the plate. It started a train of thought that eventually led to the Feynman diagrams for particle interactions and to the work that won the Nobel award.

The book swarms with delightful glimpses of the famous: Einstein, Bohr, Wheeler, Wigner, Oppenheimer, Teller, Pauli, Compton, Fermi, Gell-Mann and many others. (It is said that the California Institute of Technology, where Feynman has taught since 1950, hired Murray Gell-Mann so Feynman would have someone to talk to.) And it was, apparently, a casual remark of one of them, John von Neumann, that gave Feynman a sense of political detachment that he claims has kept him happy ever since. You don't have to feel responsible, the great mathematician told him, for the state of the world. Not for its social and political insanities, perhaps, but trying to understand how nature works on her deepest level is a responsibility — perhaps pleasure is a better word for what Feynman feels — that he surely has not taken lightly. □

Martin Gardner, formerly a staff writer for Scientific American, is a contributing editor of Discover. He is author of many books, the most recent of which is Order and Surprise (Oxford University Press, 1984).

Global matters

David R. Rosseinsky

The Cambridge Guide to the Material World.

By Rodney Cotterill.

Cambridge University Press: 1985.

Pp.352. £17.50, \$34.50.

"We are living in a material world" is the burden of a current popular song, and this inviting tome comprehensively portrays its component materials. While — despite the title — not exactly instruction against excessive monasticism (nor yet a coffee-table book, despite the large and colourful jacket) it possibly does have a message underlying its non-mathematical account of materials. Some are materials as in materials science, that is to say mainly solids (metals, polymers, ceramics and the like, together with the new organic superconductors and kindred exotica) but also included are nucleic acids, cells, living systems and their operation. More than a third of the book is devoted to biological materials, metabolism, nerve and brain operation and so on. Perhaps the philosophical drift of this mechanistic outline lies in the concluding sentence — "The nature of the mind remains a mystery". And thereby may the title be nicely justified.

The level of exposition is a mature one, even though for example the concept of potential energy itself has to be established in terms of uphill and downhill travel. Description of nuclear and atomic structures leads on to the complexity of, for instance,

the superconductivity mechanism. Experiment is included insofar as it aids the continuity of description. Lively historical exposition is often followed, and individuals are named at a pleasing narrative pace. And there are many fine illustrations, often in colour.

All is not quite perfect, however: we could do without the entropy death of the Universe (yet again), and on the atomic scale some aspects are just wrong rather than oversimplified. Thus p orbitals are not appropriately represented, the hydrogen bond and electron affinity suffer in verbal compression, the invention of the T-jump rate measurement is wrongly ascribed and so on. It is a measure of the book's overall strength, however, that the description of materials, which is the author's *métier*, survives such shortfalls in the physics and chemistry at the beginning.

What will be the readership of such a book? Certainly established physical scientists will benefit (and perhaps feel smart at spotting the occasional slip). But *all* enquiring amateur or professional scientists and technologists of whatever origin or age should enjoy it, and possibly the eye-catching jacket will fulfil the role of the Venus fly-trap depicted there. The price is such as to encourage private investment, and who knows, even as a coffee-table book it could do good by stealth. In Britain the invaluable campaign for "Women into Science and Engineering" goes under the acronym WISE, and for a WISE education, no Material Girl (of the pop song) should be without a copy. □

David R. Rosseinsky is Reader in Physical Chemistry at the University of Exeter.



Visual distraction — Maurits Escher's "Prentententoonstelling", reproduced from the chapter on crystal defects in *The Cambridge Guide to the Material World*.

New
Spring 1985

NUCLEOTIDE SEQUENCES 1985

SAVE 25%
on four-volume set

This vital reference work is the result of a unique collaboration between the EMBL Sequence Data Library staff, the staff of GenBank® at Los Alamos National Laboratory, and Bolt Beranek and Newman Inc. It contains over 4,000,000 bases from nearly 6,300 reported sequences. Keyword phrase, taxonomy, author and library location indexes assist the use of each volume.

Due April 1985. Set of four volumes: 2072pp; 0 947946 25 X (softbound); £85.00/US\$120.00.

I Primates and rodents: 524pp; 0 947946 26 8; £28.00/US\$40.00.

II Other eukaryotes: 474pp; 0 947946 27 6; £28.00/US\$40.00.

III Bacteria, bacteriophage, organelles and structural RNA: 538pp; 0 947946 28 4; £28.00/US\$40.00.

IV Viral and synthetic sequences: 536pp; 0 947946 29 2; £28.00/US\$40.00.

'... for those without computer access or for the occasional "Has this gene been sequenced?" enquiry, the printed compendium will be of great value.'

Nature (on Nucleotide Sequences 1984)

Marihuana '84

Proceedings of the Oxford Symposium on Cannabis

Edited by D J Harvey, University of Oxford; Assistant Editors: Professor Sir William Paton and G G Nahas, Columbia University and INSERM

Marihuana '84 contains reviews and papers on the latest marihuana research. The book deals with the pharmacology, metabolism and distribution of marihuana cannabinoids. It places emphasis on their adverse effects, particularly on reproductive functions and the immune response.

Papers cover the analysis of cannabinoids, focusing on the measurement of low levels of them and their metabolites in body fluids. Papers also review potentially therapeutic uses for cannabinoids, eg as anti-emetics in cancer chemotherapy.

Due May 1985; 800pp (approx); 0 904147 95 9 (hardback); £55.00/US\$99.00 (provisional)

IRL PRESS
Oxford · Washington DC

The foot inaudible and noiseless

David Park

Time's Arrows: Scientific Attitudes Toward Time.

By Richard Morris.

Simon & Schuster: 1985. Pp.240. \$17.95.

IN THE philosophies of nature that most scientists seem to work out for themselves, time is something of a dark area. Time has extension analogous to space, but the impossibility of experiencing any instant more than once defies analogies — we face great difficulties if we try to define even so simple a word as "now" without tautology. An odd fact is that all this doesn't seem to matter. To use a concept such as pH or electric potential we need to understand more or less exactly what is meant, yet we readily navigate the shoals of time even if we cannot explain what we are doing. Our sense of time passing defies explanation in rational terms (how fast does it pass, then?); but we understand perfectly what we are doing when we use a clock.

This being the case, it has often been felt that there is some need for a further explanation of time, and yet in the absence of any practical difficulties one does not quite see what needs to be explained. In *Time's Arrows* Richard Morris adopts a reasonable programme: "We must discover what science can tell us about time before we can know what the relevant philosophical questions are". His remarks on the ancients are mercifully brief. The eye picks out the phrase "the philosopher Pythagoras" and again "Ancient peoples believed that time was cyclic in character" on an early page, and prepares to close in sleep, but pushing on we arrive at Galileo and the incorporation of a time variable, for the first time, in the description of physical phenomena (in this case free fall). Now we are safely in scientific territory. I do not want to encumber this review by enumerating the author's errors, but will mention three examples of the kind of thing that a reader ought not to have to bother with in reading an otherwise illuminating discussion.

● "Galileo was the first to realize that the acceleration of these [falling] bodies was a function of time." Since this remark is made in the context of motion with constant acceleration, I can only attribute it to some demon in the pen.

● "Furthermore, Aristotle said, the velocity at which an object fell was proportional to its weight." This is indeed the conclusion that was widely drawn from passages in Aristotle (for example *Physica* 215^a and 216^a), but these make no mention of falling bodies and "Aristotle said" is hardly the way to summarize them.

● In discussing free fall Morris attributes

to Galileo the formula $d = \frac{1}{2}at^2$. Galileo never assigned a numerical value to acceleration and wrote no such formula. He did state that d varies as t^2 ; this is an immediate consequence of the fourteenth-century mean velocity rule. It would have been an interesting paragraph on the history of science if Morris had discussed the fact, surprising today, that the mediaeval scholars did not seek to apply their mathematical formulae to the natural world.

The next chapter is on "Calculus and the Idea of Determinism". After removing the sentence "Today, all of physics and most of higher mathematics are based on calculus", and placing it like a fishbone on the edge of the plate, the reader finds an excellent account of how the invention of the calculus made it possible to extrapolate present motions into the past and future so that the trajectory of a body, as a whole, could become a subject of study. The implications of Laplacian determinism are carefully discussed, though one would have hoped for at least a brief mention of the qualifications of the idea which the researches of the past decade on highly unstable motions have made necessary.

Morris next devotes two chapters to the enormous expansion of the old time-scale stemming from the work of Hutton and Lyell in geology, and Darwin and Wallace in biological evolution, and then moves on to modern physical science. The arrows of the book's title are five in number: the thermodynamic arrow of increasing entropy, the direction of cosmic expansion, the time-asymmetry of weak interactions, the electromagnetic arrow and the psychological arrow. The electromagnetic arrow, which requires that retarded solutions of the wave equation are the only ones to be retained, is the subject of an especially sensible discussion emphasizing that, in a universe containing absorbing matter, either a diverging or a converging wave can be analysed in terms of either advanced or retarded functions; the predicted phenomena are the same in either case, and so the distinction lies outside physics in the mind of the analyst. The psychological arrow involves a real puzzle: how can one understand from all we know of time why consciousness singles out a single instant, now, for its attention?

The last chapters cover the fairly familiar territory of relativity and cosmology. As a whole, this is a somewhat uneven book, serious and thoughtful in places, jejune in others, disfigured by careless errors, but finally, I think, succeeding in what the author sets out to do: to suggest to the reader some of the thoughts that come to the mind of a scientist invited to focus on the word "time". □

David Park is a Professor in the Department of Physics and Astronomy at Williams College, Williamstown, Massachusetts.

But is it art?

Andrew Dearing

Creative Computer Graphics.

By Annabel Jankel and Rocky Morton.

Cambridge University Press: 1984.

Pp.143. £15.95, \$29.95.

Art and the Computer.

By Melvin L. Prueitt.

McGraw-Hill: 1984. Pp.246. Hbk \$39.95,

£34.75; pbk \$29.95, £26.25.

"COMPUTER power" used to refer to the rate at which machines digested numbers. The users were, in the main, scientists, and they generally viewed developments in terms of a new machine's ability to solve bigger problems faster than before. Graphics devices were useful for checking results, but were rarely seen as more than that; there were thought to be better things to do with an expensive computer than use it to draw pictures. Then the price of computers fell, and we finally realized that we could control and understand the computer's calculations better when we could see the problem and the results directly. Equally, quality graphics created a market for personal and small business computers, reaching those for whom numbers held no great appeal, and allowed people working in design, manufacturing and art to discover a new medium with which to express their creativity.

Being creative involves more than producing striking imagery, however. These two books attempt in different ways to examine how computer graphics can be used in this way, and why difficulties are encountered in achieving the desired effects, while presenting a wide selection of pictures that have been produced using the medium.

Jankel and Morton survey many aspects of modern computer graphics technology in order to portray ways in which graphics

images can express and communicate ideas, or emulate the natural world. Early chapters trace the history of the discipline, and describe how an object can be modelled within the computer and how these models are converted into images on the screen. Their treatment is non-mathematical, and the novice will have no difficulty following the thread of ideas. Accompanying pictures, which throughout are of the highest quality and generally in colour, complement the text with a refreshing competence and clarity; it would have been all too easy for the authors to rely on the vividness of the pictures to compensate for an anodyne commentary.

Subsequent chapters are devoted to the ways in which computer-generated images are being applied to science, design, art and the media, and used in home computer programs and video games. The authors pay much attention to the technical and human problems of realizing the technique's potential, and there are some amusing insights along the way. One story that appears to be almost apocryphal describes how an attempt to enter a computer-generated print for a New Jersey competition was rejected in 1969, but accepted the following year when the artist presented the same piece as a hand-made silk screen. We often expect too much too soon, and technical novelty must be studied in terms of what is already known, not presented in isolation, before the artist knows how to use the medium and the viewer can understand the context within which he should see the results. Certainly, Jankel and Morton demonstrate that creativity remains the product and responsibility of the craftsman. In spite of the argument that such creativity resides in the programmed sequence of code that produces the image, the best results seem to be those in which computer graphics is used to complement and enhance material obtained using other techniques.

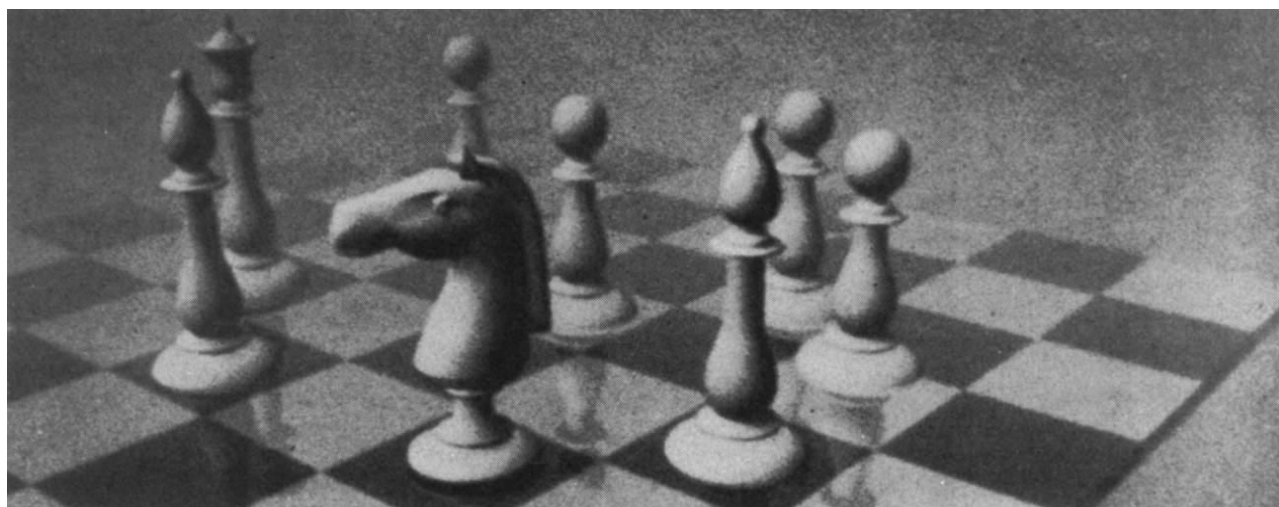
In *Art and the Computer*, Prueitt con-

cerns himself specifically with computer graphics as an art form, and most of his book is used to present a gallery of pictures representative of recent work. He too describes the methods used to generate the images, and analyses their potential as tools for the artist, but somehow neither his pictures nor the commentary really succeed in showing that computer art is yet on a par with more traditional work. Interesting topics are introduced — for example, Prueitt notes how visual appeal often depends on precise use of colour — but he fails to develop the opportunities he has created, and merely repeats his contention that the results can be considered to be art.

It is, however, not sufficient to define art simply as "creative work produced by a person", since that does no more than make us ask what the author means by "creative", and that question is not answered in the book other than by encouraging us to look at the results for ourselves. Yet, although the work reproduced here is twice claimed to be representative of the finest computer-generated art, by its greatest practitioners, over half of the pictures are by the author himself, and are mostly variations on a theme of broken symmetry, resembling pictures viewed in a kaleidoscope or drawn with a child's epicyclic toy. While they show clear technical skill, there is too little sensitivity to the need to define where the art has contributed something new.

Neither of these books offers much that is new to the specialist. Nonetheless, Jankel and Morton have succeeded in portraying the creative use of computer graphics in a way that will interest both layman and professional. Their book should remain an accurate snapshot of the state of this particular "art" in the mid-1980s. □

Andrew Dearing is currently Head of the Molecular Graphics Group at Shell Research Laboratories, Sittingbourne, Kent.



"Foggy Chessmen", by Turner Whitted and David Weimer, produced using scanline algorithms (rather than ray tracing) to produce the shaded effect. Scanline algorithms cannot cope with reflections, and the mirror chessmen are the original ones turned upside down and toned down to give the effect of reflection in a marble chessboard.

From *Creative Computer Graphics*

Imagined worlds: visions of theoretical physics

David Bohm

Imagery in Scientific Thought: Creating 20th Century Physics.

By Arthur I. Miller.

Birkhäuser: 1985. Pp. 320. \$24.95.

THE ultimate origin of creative developments in science, art and in other areas of life will probably remain a mystery, no matter how far our knowledge may progress in the future. Nevertheless, it seems reasonable to suppose that a great deal of light can be shed on the subject by a careful analysis of the process in which new ideas develop in actual historical case-studies. In *Imagery in Scientific Thought* Arthur Miller does this for the field of science, tracing the relationship of creative thinking and the construction of new scientific concepts from pre-scientific knowledge. To this end, he focuses on the work of Poincaré, Einstein, Boltzmann, Bohr and Heisenberg, because these physicists were specifically influenced in their research by their consideration of the nature of thinking itself.

In the book, Miller makes the key point that the major change that took place through the work of these scientists was the development of a fresh attitude towards mathematics and the imagination. Thus, Poincaré's view on the subject was that Euclidean geometry arose necessarily from man's earliest thought because it was needed for his survival. For Poincaré this implied the basically conventional nature of the axioms of geometry. To obtain real content, they have to be enriched with lower-level physical hypotheses. The introduction of such hypotheses is through *invention*, and this comes out of an intuition that cannot be defined explicitly but must be felt. This is a kind of Baconian inductivism with gaps supplied by intuitively guided inventions. Physical theories then express in the most convenient manner the "true relations between real objects" along with the invariant laws pertaining in these relationships.

Einstein adopted much of Poincaré's approach, and explicitly expressed his debt to Poincaré in calling attention to the need to "pay strict attention to the relation of reality and our concepts". However, he rejected Poincaré's *a priori* organizing principles along with his closely related assumption of a fundamental role in our thinking for Euclidean concepts. Rather, he regarded concepts in general as admissible "only to the extent that observable facts can be assigned to them without ambiguity". Thus, for Einstein, the question of whether Euclidean geometry is valid was basically *physical* (for example, that

of whether it is possible to attach the Euclidean concept of length to a quasi-rigid measuring rod in an unambiguous way).

Einstein considered science to be essentially a refinement of everyday thinking, in which images are of course involved. He felt it necessary to analyse serious thinking (in contrast to "dreaming") as a series of such images ordered by a recurrent image that constantly turns up. Such an ordering element can relate many different series of hitherto unconnected elements, and this ordering element he regarded as a concept.

Though he seems to have started from Poincaré's approach of building a theory step by step by posing hypotheses to explain the experimental data, Einstein realized by the end of 1904 that the "known facts" were too restricted to solve the problem of an electromagnetic world picture. Instead he shifted to an appeal to general axioms and theories of principle, similar in structure to those of thermodynamics. These were the principle of relativity of motion and the principle of the invariance of the speed of light. Here, Einstein's thought was affected by the views of Boltzmann. For Boltzmann, our mental pictures, which he regarded as abstractions from objects actually perceived, must be clearly and unambiguously imaginable. A striking example of this is Einstein's highly imaginative picture of an observer who moves along with a light wave and sees only a static configuration of fields. The implication of this picture violated Einstein's physical intuition that the electromagnetic field is inherently dynamic, and eventually led to his taking the invariance of the speed of light as a basic principle.

Through the consideration of the properties of dynamos (as well as in other ways) Einstein was then led to regard Lorentz's idea of a local time in terms of which Maxwell's equations were invariant (to first order in v/c) as similarly indicating a general principle, which was the relativity of simultaneity (and which held to all orders in v/c). Here, Einstein was clearly guided by the notion that our ordinary concepts of time can no longer be related unambiguously to the experimental facts when we consider the full implications of electrodynamics (as he showed with the aid of various hypothetical experiments). So he was ready to go against ordinary intuitive experience when he felt this to be necessary, but he still insisted that concepts have to be *some* kind of abstractions from our pre-scientific perceptual experience which must be unambiguous in the senses indicated above.

This brings us to the contributions of Bohr and Heisenberg. Their analyses of the meaning of the quantum theory indicated to them that there are limitations on the possibility of unambiguous intuitive visualizability of the properties of matter at the quantum level. Bohr emphasized the need to separate the space-time and the causal descriptions. This implied that the

ordinary imaginative pictures can be retained, but that they have to be *ambiguous* related to experimental facts. Here, perhaps, is the fundamental difference between Bohr and Einstein, who, as we recall, insisted on the need for the possibility of an *unambiguous* relationship between concepts and the experimental facts. In my opinion, it is unfortunate that Miller does not bring this point out in the book, as it is of crucial importance in understanding the split that eventually came about between Bohr and Einstein, a split that profoundly affected the development of quantum theory and relativity.

Heisenberg's initial approach was, in contrast to that of Bohr, to insist that our customary intuitive visualization (*Anschauung*) could not be extended into the atomic domain at all. Miller follows the further development of Heisenberg's approach in some detail, to show how it led eventually to a new notion of intuition and imagination, which Heisenberg called *Anschaulichkeit*. This referred not to ordinary visualization, but rather to the imaginative display of the meaning of the mathematical formalism (which latter was by then regarded as the most direct and fundamental way possible of putting the truth about nature). This mode of visualization is now most commonly encountered in the Feynman diagrams which do not correspond to reflections of actual physical processes, but which serve primarily to enable physicists to think imaginatively about the meaning of the mathematics. As Heisenberg himself said, this is essentially an extension of the Platonic (and perhaps Pythagorean) approach. It may be summed up in the dictum of Jeans: "God is a mathematician".

If one looks at how modern theoretical physics is actually done, one can see that Heisenberg's approach now dominates, not only in particle physics but also in cosmology. Miller's book provides important insight into how the whole way of doing theoretical physics thus altered radically in a comparatively short time. This insight could, however, perhaps have been extended if the author had brought out, even if only briefly, how several centuries ago a similar change occurred, culminating in the Newtonian approach. To have done this would have indicated that what constitutes commonly accepted scientific practice has constantly been changing, and that it may therefore change again in an equally fundamental way. Tacitly to regard the currently accepted way of doing physics as the last word on the subject is to encourage the same sort of absolutist conclusions that took hold of the Newtonian way of doing physics, and thus help to ossify the current approach too.

The latter part of the book is a survey of various attempts of modern psychologists to account for the creative development of modern physics. It begins with a description of Wertheimer's application of Gestalt

psychology to Einstein's process of thinking in the work that led to the theory of relativity (Wertheimer's work drew upon actual discussions with Einstein himself). Miller then goes on to a broader discussion in terms of new researches based on extensions of Piaget's work on genetic epistemology. Two different versions of genetic epistemology are considered here, the first due to Pylyshin and the second to Kosslyn. Pylyshin and Kosslyn are in agreement in that they develop detailed accounts of how imaginative thought actually proceeds and try to support these accounts on the basis of actual case-studies. Their principal disagreement is on the dynamic role of images. Roughly speaking, Kosslyn assigns a positive predictive power to images, while Pylyshin regards them, in effect, as little more than passive displays of abstract propositional thought.

Although Miller gives what seems to be a fair treatment of this whole subject, I feel that he fails to be adequately critical in that he accepts an assumption that seems to be taken for granted in all the work on genetic epistemology. This is that creativity is primarily to be understood through the development of knowledge, enriched perhaps from time to time by *invention* whose source is (along the lines indicated by Poincaré) ultimately in a mysterious intuition. I would like to suggest that while all of these factors are involved in creativity, its primary source is in *perception*, not only through the senses, but, even more, through the mind (as indicated by the Greek *noesis* and the German *Vernunft*, which latter means intuitive or perceptive reason). Kostler gives an example of this, in the sudden realization of Archimedes that the volume of water displaced by a body is the same as the volume of the body independently of the complex details of its shape. Perhaps we could say that Einstein must have similarly realized somehow that the problem with electrodynamics was essentially the same as with thermodynamics; that is, to find principles that are independent of a detailed dynamical description of all the complex processes that are actually at work in concrete experimental situations. It seems to me that a study of the history of the subject from such a perspective is necessary, if we are to have a proper account of how the mind gets out of the "rut" of thinking in terms of old concepts that have long been taken for granted.

I feel that in spite of certain inadequacies Professor Miller has made a valuable contribution to our understanding of creativity in twentieth-century physics. His book is well worth reading by all those who are interested in the history and philosophy of science, or in the understanding of the nature of creativity and its relationship to the process of thought. □

David Bohm is Emeritus Professor of Theoretical Physics at Birkbeck College, University of London.

From Cambridge to Moscow

Nevill Mott

Kapitza, Rutherford, and the Kremlin.

By Lawrence Badash.

Yale University Press: 1985. Pp.129. \$20, £20.

LAWRENCE Badash has dedicated this book to the memories of Ernest Rutherford (1871–1937) and Peter Kapitza (1894–1984). The story of the relationship between the two men is well-known. Kapitza, a brilliant young pupil of A.F. Joffe in Leningrad, lost his wife and two children in the influenza epidemic of 1921, and, in an effort to rescue him from his deep grief, friends arranged a trip abroad. They were successful; Kapitza knew where he wanted to go and arrived that year in Rutherford's Cavendish Laboratory. The Mond Laboratory was built for him and a Royal Society chair became his. There he developed equipment for producing strong magnetic fields and low temperatures, work on which was to culminate in the award of a Nobel Prize in 1978. The friendship between the two men must have been very deep, particularly as these branches of physics were far away from Rutherford's beloved nucleus.

Kapitza, as I remember, used to boast that he was the only Soviet citizen whose passport was endorsed for unlimited travel in and out of the Soviet Union. In 1934 he went with his second wife Anna to the congress in Leningrad which celebrated the one-hundredth anniversary of the birth of Mendeleef. They took their car to Bergen, and motored through Norway and Finland. I remember this, as I was invited to the conference, probably because I had helped Frenkel with the English version of his book on liquids, and Kapitza suggested that my wife and I should join them in their car (in the event there seemed not to be enough room for us and the luggage, and we went by boat). Once there, Kapitza was told that his country needed him and he must stay. Anna went back to Cambridge, joining him with the children more than a year later.

In the book Badash tells the story of Rutherford's efforts to help him, reproduces many of Kapitza's agonized letters to his wife and to Rutherford, and records how eventually much of the Cambridge equipment was bought by the Soviet government, and how, two years later, he was able to settle down to work in his new Institute for Physical Problems of the Academy of Science. He had the help of two of his technicians from Cambridge, who spent several years with him in Moscow.

Kapitza's misery in his first year, which might have seriously endangered his health, was not because he was unwilling to work for his country; he always protested that he wanted to. It was because, until the Institute was ready, there was nothing he could do. To be cut off from work was dreadful for him. If he could have returned to Cambridge, supervised the transfer of equipment to Moscow, finished off certain researches and indeed been treated like a loyal Soviet citizen, he would not have complained, or at any rate not so bitterly. But — and this is what made life intolerable — he was given no trust. People were afraid to talk to him and so he had few friends. An official was detailed to look after him, and promised much, but when Kapitza's winter clothes were posted to Moscow from Cambridge he could not even get the parcel through customs without Kapitza having to pay a huge import duty. Kapitza was denied Western newspapers, but letters to and from his wife seemed to get through, as did telephone calls. I was once in R.H. Fowler's house when a call came through from Kapitza, which lasted a very long time. Apparently he could at least do that, and trunk calls must have been cheap.

A letter from Kapitza to Rutherford in the early days of the Institute records vividly the difficulties of doing research at that time under the Soviet system:

You see Soviet industry is growing at a terrific rate and all is done to make its growing organised and well planned, so that all the system of supply and production is also well planned and organised. But the supply of the factory, which is operating according to a definite plan, must be anticipated in details at the beginning of the year, and evidently represents its demands in very large figures. Such a system is indeed quite unsuitable for the supply of laboratories. I have written and spoken to the



From Kapitza, Rutherford, and the Kremlin

Peter Kapitza (about 1925) and the Institute for Physical Problems, Moscow.

authorities that the labs must be supplied in a different way, and I think that the people here begin to recognise that the system of supply must be altered for the scientific institutions. But, at present, if for instance we require four bars of phosphor bronze, the need of which we had not anticipated in the beginning of the year, we must get it as an exception and the permission of the sub-secretary of the Heavy Industries is required and this means a lot of correspondence, which is equally extensive if we require 10 kilograms, 10 tons or 10 loads of stuff. And so even if the Institute is given sufficient money, and even if the industries are quite decent, we are actually very badly supplied.

He believed, however, that things would get better. But later he wrote:

My colleagues the scientists are very scared of me and behave like pigs. My institute is attached to the Academy of Science, of which I am not a member but they govern my institute. The president is Karpinsky, he is ninety years old. During meetings of the Council he sleeps with a happy and kind smile on his face.

It is for insights such as these, showing the reactions to Soviet bureaucracy of a

great man who had experienced Cambridge, that I recommend this book. Parts of it are compulsive reading. How much the bureaucracy has changed since then is not discussed but in any case Kapitza came to terms with it, became a Hero of Socialist Labour, received several Orders of Lenin and was throughout a sane and courageous voice in Soviet science. One wonders if he would have survived the early days without Rutherford's moral support, and Rutherford's efforts — though unsuccessful — to secure his release, which are here described by Badash. Kapitza repaid some of his debt in his "Recollections of Lord Rutherford", published in the *Proceedings of the Royal Society* in 1966.

Kapitza, then, was back to work by 1936. The author remarks that the timing was fortunate, for Rutherford, upon whom he depended so much for aid and encouragement, died in October 1937. □

Sir Nevill Mott is Emeritus Professor of Physics at the University of Cambridge.

Between machination and merit

Robert Fox

The Beginnings of the Nobel Institution: The Science Prizes, 1901–1915.

By Elisabeth Crawford.
Cambridge University Press: 1985.
Pp. 281. £22.50, \$34.50.

A FEW days after Alfred Nobel's death in December 1896, his close relatives gathered expectantly for the reading of his will. They were in for a shock. For Nobel had bequeathed to his family a mere one million crowns. The rest of his vast estate, amounting to 30 times that amount (over £1½ million), was to be invested to provide five annual prizes, for physics, chemistry, medicine or physiology, literature, and what Nobel (a convinced pacifist, despite the fortune he had made from explosives) called "contributions to the fraternity between nations". Predictably, the will was contested. But the adjustment that was made in favour of the family was meagre, and in 1901 the first prizes were awarded essentially in accordance with Nobel's wishes.

Elisabeth Crawford traces the difficult birth of the Nobel system and the labyrinthine processes by which the awards were made in physics and chemistry between 1901 and 1915. The sheer size of the prizes, each of them worth about 30 times the annual salary of a university professor, was enough to ensure intense public interest. And the element of spectacle was further enhanced by national as well as personal rivalries which made it difficult for the brash newcomer to integrate with the rather chaste reward system of European

science as it existed at the beginning of the century. Tradition demanded, of course, that the cruder rivalries should be concealed. But Crawford penetrates this facade to excellent effect. Drawing on her intimate knowledge of the relevant Nobel archives (open since 1974, in accordance with a new 50-year rule) and of the context and content of Swedish science in the early twentieth century, she casts radically new light on the complex interaction between the dispersed international networks of nominators and the subject committees in Sweden.

Few readers will be surprised by Crawford's demonstration of the keen opposition between the various national traditions in science, for, by 1900, science was emerging as a battle-field on which the European powers were only too ready to try their strength, not least with the disturbing new force in research, the United States. The French, it seems, were especially prone to nominate their fellow-countrymen, far more so than the Germans, who have a commendably low score on Crawford's ingenious index of chauvinism. (Britain's record was mixed, with the physicists appearing distinctly more chauvinistic than the chemists.) However, Crawford's evidence suggests that it was the Swedes themselves who injected the greatest element of competitiveness. Here, the main antagonists were Svante Arrhenius, already distinguished internationally for his work on the theory of solutions, and the mathematician Gösta Mittag-Leffler. The two men were poles apart in their political sympathies and in temperament and life-style. Arrhenius was liberal, popular and modest to the point of insecurity about his scientific abilities, whereas Mittag-Leffler was conservative, flamboyant and more respected than liked. As Crawford shows,

Arrhenius and Mittag-Leffler intervened rather irregularly, but whenever they felt moved to support a personal preference, or to thwart the candidate of a rival, they acted with rare zest.

In this running confrontation, Arrhenius had a notable early success when he saw to it that the first prize in chemistry went to J.H. van't Hoff, whose work on electrolytic phenomena complemented his

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Arrhenius — "modest to the point of insecurity about his scientific abilities . . ."

own. In 1909, Arrhenius successfully backed another ionist, Wilhelm Ostwald, and, using more negative tactics, he managed to delay an award for his rival Walther Nernst until 1921 (15 years after the first nominations for Nernst had been received). Mittag-Leffler, by contrast, saw his main mission as the advancement of mathematical physics. It was he who secured at least a share in the physics prize for H.A. Lorentz in 1902. And it was only misfortune which denied him his greatest triumph: his long campaign in favour of Poincaré had still not succeeded when Poincaré died in 1912.

These episodes and the intolerance they imply are far from edifying. But generally Crawford's book is reassuring. It provides ample evidence that although animosity, national interest and the stage-managing of nominations all played their part, scientific merit was nevertheless consistently rewarded. It is arguable that a blunder was made in 1908, when the French experimentalist Gabriel Lippmann emerged as a successful compromise candidate, following Mittag-Leffler's mishandling of his own campaign in favour of a joint award to Planck and Wien. But while Lippmann was no Planck, even this decision rewarded excellence of a kind. Like the overwhelming majority of the decisions discussed by Crawford, it seems a beacon of wisdom when compared, say, with the award of the first prize in literature to the poet Sully Prudhomme as a sop to French pride. All told, the physicists and chemists of Europe and North America come out of this intriguing and well-told story with a good deal of credit. □

Robert Fox is British Academy Reader in the Humanities at the University of Lancaster.

On trans-scientific turf

Owen Gingerich

Beyond Velikovsky: The History of a Public Controversy.

By Henry H. Bauer.

University of Illinois Press/Harper & Row: 1985. Pp.354. \$21.95, £24.25.

"THE thousands of words spilled to refute and condemn his theories gave them a sort of status, even if only that of divinely inspired nonsense." So the London *Times* summed up the controversy surrounding the work of Dr Immanuel Velikovsky when he died in November 1979.

Basically, Velikovsky claimed that in the ages before Christ a large body had been ejected from Jupiter and had gone into orbit around the Sun, eventually to become the planet Venus. On its way, it came close enough to the Earth around 1500 BC to cause the pharonic plagues of *Exodus* and may other miraculous phenomena recorded in the Old Testament, including the day the Sun stood still at Gibeon. Relying on eclectic interpretations of world-wide mythology, Velikovsky amazed his readers with his erudition and dismayed astronomers by his flagrant disregard for Newtonian mechanics.

The hue and cry against Velikovsky's theory came primarily in two waves. When his principal work, *Worlds in Collision*, appeared in 1950, heralded by sensational abridgements in *Harper's*, *Colliers* and the *Reader's Digest*, astronomers stirred up such a tempest that Velikovsky's place in the best-seller list was assured. Having learned that any publicity is good publicity, the scientists quickly lapsed into indignant silence, hoping that the interest in Velikovsky's theories would atrophy under their careful lack of attention. Their scheme may well have worked except that Velikovsky was "rediscovered" by the anti-establishment generation of the 1960s, for whom he became something of a folk hero. The scornful attacks of the early 1950s, and especially tales of a threatened textbook boycott against his publisher, Macmillan, helped build up an image of a scientific martyr, a modern-day Galileo alternately jeered and ignored by the entrenched scientific establishment.

Thus, by 1974, several of us agreed that a forum should be provided for Velikovsky to meet his scientific critics, and for the scientists to put on the record a clear statement of their objections. The actual encounter, at a meeting of the American Association for the Advancement of Science in San Francisco, proved to be a memorable non-meeting of minds, with a highly polarized audience full of partisans unlikely to be swayed by arguments or counter-arguments from either opposing camp. Scientists who had advocated silence

on the matter went away with an "I told you so" attitude. Nevertheless, this second wave of criticism was harder for Velikovsky's supporters to cope with, and Velikovsky felt particularly stung by the witty and articulate attack made by Carl Sagan. In colourful terms, Sagan described such things as the "fly ablation problem", that is, the difficulty of flies originating on the Venus-comet in entering our atmosphere just in time to cause the pharonic plague of pests.

Today, although Velikovsky's catastrophism is rapidly being forgotten, the fierce controversy that swirled for three decades can still teach us something about the nature of science and trans-scientific

Unreal Science

SIR,—As a qualified and practising scientist I wish to dissociate myself from your editorial of April 12. You allude to Geller and Velikovsky with terms like "nonsenses", "beliefs beyond science", and "pseudo-scientific ideas", and contrast these with "real science" and "science based on conventional ideas about the way a scientific investigation should proceed", and you declare that "scientists want to fight this distressing drift from the scientific way of thinking".

History is littered with ideas shown to be false by people bold enough to question their contemporary conventional science, often in the face of personal ridicule and even persecution.

Presumably you class Einstein's

Passions run high — a letter in Nature, 28 June 1974, in response to an editorial on Velikovsky.

challenges. It is such questions that Henry H. Bauer, Dean of the College of Arts and Sciences at the Virginia Polytechnic Institute, has explored. A personal and somewhat idiosyncratic account, *Beyond Velikovsky* contains material to infuriate or at least exasperate everybody. Nevertheless, it is the most balanced and informative analysis yet to emerge from the whole affair.

Bauer argues that science does not find absolute truth, but only a high degree of probability. This prevents scientists from claiming that any particular sequence (such as Velikovsky's) is impossible, only that it is staggeringly unlikely. However, Velikovsky never claimed that his scenario was probable, only that it happened. Hence, counter-arguments that it was improbable completely missed the point. Velikovsky was staking his claims on an entirely different turf — "trans-scientific" in Bauer's terminology — and so the normal type of scientific argument lost its force against Velikovsky and his supporters. In their frustration, scientists were frequently reduced to appeals to authority (their own, of course), which went directly counter to their own claims of how science worked.

Velikovsky, too, was frustrated by the resistance from the scientific community, so he consulted his neighbour at Princeton, Albert Einstein, about the status of his theory. Einstein assured him that it would be better if the theory could *predict* something. Thereafter Velikovsky searched his writings and came up with a series of "advance claims" that he had made, ranging from high temperatures on Venus to radio noise from Jupiter. Unfortunately, Einstein had not made it clear to him (and thus to Velikovsky's followers) that predictions have to be made for the right sorts of reasons to be persuasive. The principles must fit with the coherence of scientific understanding, and they should not be single *ad hoc* instances incapable of supporting further, related predictions.

Bauer analyses both Velikovsky's physical reasoning and his advance claims, and both are found wanting. Although Velikovsky's scenario cannot be denied outright, it is relatively easy to show that his understanding of physical principles was hopelessly faulty. "There is rarely a certain way of identifying a crank", Bauer cautiously asserts, and only after 150 pages does he label Velikovsky one in the sense of a person "enthusiastically possessed by a particular crotchet or hobby". But in a devastating chapter entitled "A Nontechnical Case against Velikovsky" he leaves no doubt that the enthusiastic Dr Velikovsky was also a crank in the more pejorative sense of the word.

Nor do Velikovsky's opponents escape Bauer's criticism. "In retrospect", he writes, "it seems clear that Velikovsky's critics did not grasp the magnitude — or perhaps even the nature — of the task of discrediting Velikovsky in the eyes of the public at large". Bauer charges them with errors of commission — arguing *ex cathedra*, with ridicule, *ad hominem* — as well as with errors of omission. In the latter category is, primarily, the failure to capitalize on the unoriginal character of Velikovsky's scheme.

There is, however, one logical path that, curiously enough, Bauer has not explored. Although science cannot prove that a Velikovskian scenario is impossible, it might well prove that it did not happen. At the San Francisco confrontation, Peter Huber provided evidence from Babylonian records that Venus in 1600 BC behaved just as it did in the centuries immediately before our own era. I remarked at the time that if Huber was right (as I certainly believe), Velikovsky's scenario was finished. Velikovsky tossed one of his *ad hominem* barbs at Huber, labelling him a self-confessed amateur in Babylonian. Huber, who is one of the world's leading statisticians and a redoubtable cuneiformist on the side, quickly proved on the stage how ignorant Velikovsky himself was of cuneiform languages. The Velikovskites were naturally reluctant to admit defeat by a single stroke, so they have since worked hard to discredit the Babylonian data,

which were unavoidably contaminated with copyists' errors (and hence Huber's statistical approach). Ambiguous data often characterize the cutting edge of science, and discussion of the point would have added another dimension to Bauer's analysis.

With the decline of interest in Velikovsky, another, surprisingly similar, challenge is facing the scientific establishment: so-called "Creation Science" whose extreme form calls for a 6,000-year-old Earth geologically shaped by Noah's flood. Fossil footprints of a giant humanoid beside indisputably genuine dinosaur prints in the rocks of the Paluxy Creek, Texas, are taken quite seriously by many of these fundamentalist, Biblically-literalistic people. Bauer's closing chapters on "Means of Persuasion" and "Some Realities about Science" offer sound advice about pitfalls to be avoided in these "trans-scientific" debates, where probabilities, values and desirability are at stake. How successfully his analysis can be brought to

bear upon this new challenge remains to be seen. But arguing that the creationist ideas have been heard before does not seem to be a cogent strategy. Rather, we must argue that even if science is "never completely final, always subject to question and doubt" (as Einstein put it), nevertheless it is of a piece, not easily shredded, and the process of building this coherent tapestry is what science is all about.

Beyond Velikovsky is awkwardly repetitive, not as complete as it might appear at first glance and in its own way rather too polemical. Its system of references, numbered but unattributed in the text itself, is seriously distracting. Nevertheless, the account is detailed, thought provoking, and the best yet written on a bizarre episode that continues to smoulder in the annals of science. □

Owen Gingerich is Professor of Astronomy and the History of Science at the Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts.

Continental epic

Robert M. Hazen

Landprints: On the Magnificent American Landscape.

By Walter Sullivan.

Times Books, New York: 1984. Pp.384. \$22.50.

"THERE is no epic tale to compare with the account of how the North American continent came to be." So begins *Landprints*, one of the finest popular earth science books to appear in decades. In this recounting of the 3,500-million-year evolution of the mountains and valleys, rivers and deserts of North America, Walter Sullivan, the science editor of the *New York Times*, has achieved a masterly blend of geomorphology, historical geology, geophysics, structural geology, palaeontology and geodynamics.

The book consists of 37 short, self-contained essays on the principal features of the American landscape. Inspiration for this geological feast came from the author's frequent transcontinental flights; as Sullivan observes, many of the continent's landforms are best viewed from the air. The first two dozen chapters, therefore, focus on striking geomorphological features — parallel ridge systems, meandering rivers, volcanic craters and cones. The White Mountains, Wyoming basins, Hawaiian Islands and the Grenville Province are among the many areas which are examined first in terms of their distinctive aerial appearance, but then also from the standpoint of their underlying geological structure and history. In this way Sullivan gradually introduces current views on plate tectonics, continental subduction, and the regional folding and faulting that

are associated with mountain building.

The last third of the book emphasizes the dramatic alterations of the Earth's surface imposed by man. Farms and mines, cities and suburbia have all left their distinctive imprint on North America, and Sullivan discusses both their geological and environmental significance. He explains why East Coast cities grew up along a former continental shelf margin (the "fall line"); he shows how rock erosion influences soil chemistry, which in turn controls the predominant crops of American farms. From aerial observations of the distinctive patterns caused by centre-pivot irrigation, hydroelectric dams and strip mining, Sullivan proceeds finally to consider the possible consequences of man's actions on his environment.

Landprints is graced by a magnificent collection of photographs and other illustrations. Aerial and Landsat photos — more than 150 of them — display the variety and splendour of North American topography, both natural and man-made. Views of ancient craters, barrier islands, dune fields and the crazy-quilt of America's farms, photographs of road-cuts, regional maps and schematic diagrams of geological processes, all enrich the presentation.

This is a beautifully produced book, too, with clear typography, a felicitous layout and sturdy binding. The writing is stylish and informative, in the best tradition of popular scientific literature; Sullivan brings to his subject an infectious sense of wonder and enthusiasm — he entertains as he teaches. Anyone interested in science, particularly those who travel frequently by air, will be well rewarded in reading this continental epic. □

Robert M. Hazen is at the Carnegie Institution of Washington's Geophysical Laboratory, Washington, DC.

Advertisement



The natural history of the primates

J.R. & P.H. Napier

A superb introduction to the most fascinating group of animals on earth. Desmond Morris

This authoritative and detailed account describes the behaviour and evolution of all 57 genera of primates living today. Comprehensive and highly readable it will be a valued source of information for both amateur naturalists and professional zoologists and researchers. 200pp, lavish colour illustrations. 0 565 00870 6. Hardback £15.00



Ferns of Jamaica

G.R. Proctor

The definitive nature of this taxonomic record and description of the 579 species and 30 varieties of ferns occurring in Jamaica will ensure its status as a reference work for many years to come. All genera are illustrated and full distributional data is given. 631pp, 135 line drawings, 22 maps. 0 565 00895 1. Hardback £50.00

Bats: a natural history

John E. Hill & James D. Smith

Well written and easily assimilated . . . a great deal of hitherto unpublished material . . . excellent line diagrams . . . a very useful book P.A. Racey - Nature
After selling out in its first year this popular book has now been reprinted. 248 pp, illustrated throughout. 0 565 00877 3. Hardback £15.00

Staying alive - Computer software

This visually exciting and instructive program is based on one used by thousands of visitors in the Museum's *Hall of Human Biology*. It provides a practical demonstration of the principles of homeostasis and is especially relevant to O-level and CSE Biology and Human Biology curricula. Cassette for BBC Model B Micro-computer. Includes instructions. 4 colour moving visual display. 0 565 00987 7. £7.00 inc. VAT

Publications Sales, British Museum (Natural History), Cromwell Road, London SW7 5BD.

Earth science in a fabulous place

Peter J. Smith

Atlantis in the Light of Modern Research.

By Zdeněk Kukal.

Elsevier: 1984. Pp.224. \$57.75, Dfl. 150.

How the childhood memories come flooding back! The idea of Atlantis was introduced early into my genteel upbringing, presumably as a spur to the imagination much as a spot of breakfast is supposed to set the digestive processes rolling. Plato's lost world beat its competitor, Little Noddy, hands down, losing its status as prime imaginative wonder of the world only under the impact of the higher Blytonian *Island of Adventure* and, later, the more robust yarns of Percy F. Westerman and Arthur Conan Doyle. Besides, in the black-and-white world of the teens, fiction was fiction and fact was fact, but fiction masquerading as fact (as, by then, we all knew Atlantis to be) was strictly for the gullible. When we became adults we put away childish things, did we not?

But no, actually, some of us didn't. N.F. Zhironov of the Soviet Union, reputedly the world's most devoted "atlantologist", is even now engaged in a review of all of the estimated 3,600 works on Atlantis published to date. A Dr E. Sykes is said to edit *Atlantean Research*, a periodical containing such gems as "Atlantis, land of giants" and "Atlantian script". Pierre Benoit wrote the most famous of all novels about the place. And Zdeněk Kukal, a member of the Geological Survey of Czechoslovakia and author of more than 200 legitimate papers and books on matters geological, now presents the results of 20 years' study of the scientific aspects of Atlantis.

What these four representatives of a much larger motley demonstrate, of course, is that Atlantis has a number of distinct faces. There is Atlantis as jumping-off point for harmless fantasy, which sees Plato as an early science-fiction writer. To the extent that fantasy is therapeutic, here surely lies Atlantis's most abiding contribution to the welfare of the human race, although Kukal ignores all that because it is not science. Then there is Atlantis as advanced civilization. But is it likely that there was really, as Plato claimed, a ten-million-strong community that could write, work metals and breed domestic animals many millennia before the time for which there is evidence of such things elsewhere? Kukal thinks not.

That leaves Atlantis as a physical, geographical and geological entity, albeit one that, according to Plato, suffered "violent earthquakes and floods, and in a single day and night of misfortune . . . disappeared

in the depth of the sea" about 11,500 years ago. Not that Plato's sea should be taken too literally, it seems, for of the 40-odd possible locations of Atlantis given in the literature (ranging from the United States to the Black Sea and from Iceland to Zimbabwe) almost half are on land. As far as the ocean floor is concerned, however, Kukal has to admit that echo sounding, deep-sea drilling, dredge and grab sampling, underwater photography and bathymetric descents have revealed not one whiff of Plato's place.

From this discouraging beginning Kukal goes on to assess in some detail whether we could just have missed the sunken state (possibly), whether city walls and metal artefacts could have survived 11,500 years of submergence anyway (highly unlikely), whether Atlantis could have been any of the fictitious islands shown on early maps (no), whether it could have been any of today's real islands (geologically impossible) and whether there are any areas of subsidence in the Atlantic or Mediterranean which might qualify (not really). He then considers whether the supposed Atlantis could have been destroyed by earthquake or tsunami (theoretically possible), storm flood (probably not), volcanic eruption (yes), slow tectonic movement (no), cosmic impact (no evidence) or rising sea level (yes, but not "in a single day and night").

There is something faintly ludicrous about all this; it is a bit like mounting a scientific search for the invisible man simply because H.G. Wells happened to write about one. On the other hand, in describing the search for Atlantis, Kukal does manage to impart a lot of earth science — the geology of islands and oceans, the nature of glaciation, the role of sea-level changes, the processes of weathering and much else besides. The means may not justify the ends, but they make a fascinating story all the same.

I began the book with an inclination to mock, but ended up with a grudging respect for Kukal. He has produced a fine and objective refutation of a "theory" that may be inherently ridiculous but is nevertheless important because it has duped thousands of people over the centuries. Atlantis was a deliberate fiction and a particularly successful one at that, for it still attracts attention after 2,500 years. The advanced civilization of Atlantis was Plato's Utopia; its physical and political structure was the best Plato could envisage; and its destruction was by processes known to him. There can be no other conclusion, although one almost wishes that after 20 years of work Kukal could have proved otherwise. Of course, the only people to accept his verdict will be those who believed it in the first place. Those who don't never will. All will continue much as if Kukal had never been. □

Peter J. Smith is Reader in Earth Sciences at the Open University.

Swimming through the cell

Tim Hunt

A Guided Tour of the Living Cell, Vols 1 and 2.

By Christian de Duve.

W.H. Freeman: 1985. Pp.423. \$55.95, £39.90 per set.

HERE is the latest biological work from W.H. Freeman in the Scientific American Library. It spans two volumes, but is a seamless whole and the two jackets are identical except for the words "One" and "Two". Like its predecessors in the series, *A Guided Tour of the Living Cell* is beautifully produced and written, and its author uniquely qualified to show us around. Christian de Duve is one of the great pioneers of cell biology; together with Albert Claude and George Palade, he was justly awarded the Nobel Prize in medicine in 1974 for his "discoveries concerning the structural and functional organization of the cell".

You may remember a film called *The Incredible Voyage* in which heroic brain surgeons operated on the patient by shrinking themselves down to a microscopic size and making their way to the site of the damage in a tiny submarine which was injected into a vein. The operation successfully completed, they escaped by means of the tear duct after many a frightening adventure, the worst an ugly encounter with a macrophage. In his books de Duve shrinks the reader even further, to the point where the very organelles of the cell are as huge caverns to the observer. What is more, we are protected only by a diving suit (a magic suit, it's true, made of the material that coats the leprosy bacillus) as we swim around the cell with the Master as our guide. The idea sounds perilously twee, and when the introduction reveals that these are the books of a series of childrens' Christmas lectures, the heart sinks even further. However, my initial suspicion wore off within a few pages and the idea works brilliantly well — they must have been fabulous lectures!

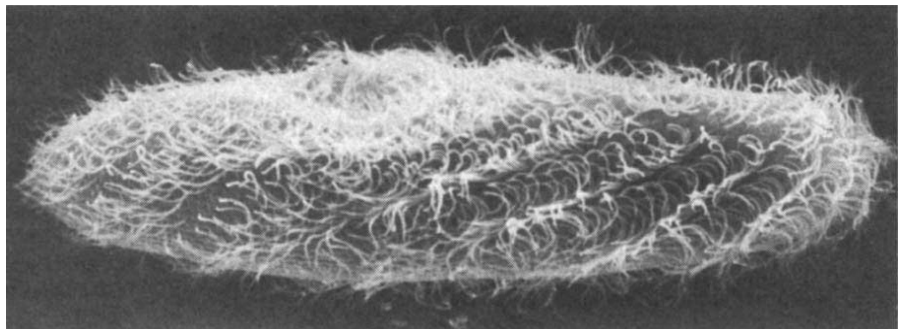
Not surprisingly, de Duve is at his best describing the workings of the membrane systems of the cell, and particularly in describing lysosomes. As he says "understanding, not description, is the true purpose of scientific exploration", and in this chapter understanding comes as if by enchantment. We are lucky in other ways besides. de Duve's language is plain, his images concrete and familiar, and the text is strewn with wise words. Several times I found myself warming, seeing things said that do not often appear in books of any kind, let alone relatively elementary texts. Other agreeable features of the general style are de Duve's readiness to admit ignorance and to point out

puzzles, and his unfailing grace in explaining the (usually classical) origins of terms.

However, the *Incredible Voyage* approach has its limitations, most evident in the biochemical parts of the books. Very few people really enjoy biochemistry, probably because of its rather abstract nature. It's a subject that's fun to *do*, but not so much fun to read about. de Duve adopts two "outlandish terms" (sic) the "Janus Intermediate" and the "Oxphos Unit". The Janus Intermediate "arises from a nucleophilic attack by an oxygen containing building block (X-OH or X-O) on

exploring at greater length, precisely because it seems to cause such trouble. de Duve includes some appendices dealing with biochemistry and bioenergetics which would have been helpful in the body of the text, but they are at the end of the second volume so consulting them is not encouraged. Probably you either have to grasp the nettle and give the full treatment, or leave it out altogether.

In the nucleus we are for the most part in conceptual country which de Duve understands very well, but doesn't know like he knows his organelles. de Duve is not alone in speaking of "the nucleus" as stan-



Stopped in time — rows of cilia on the surface of *Paramecium tetraurelia* arrested in the middle of their synchronous beat.

ATP or some related energy-rich molecule, which we will designate provisionally as A-B, in order to avoid going into complex chemistry"; later we learn that "Janus...is Mercury and Aquarius all in one". You see? The excellent colour graphics make it more palatable, but whereas other parts of the book repeatedly opened my eyes, I found the biochemistry and bioenergetics not only unilluminating but even baffling in places. I never really did get the Oxphos unit.

In his introduction, de Duve is apologetic about his approach to biochemistry, as if he knew he was taking a bit of a risk and hadn't quite got it the way he would have liked. It would be interesting to know how a non-biochemical reader would fare. It is very difficult to know how to deal with biochemistry in a modern account of the cell. Whether you are a strict cell biologist or a full-blooded clone-and-sequence molecular man, enzymes and the reactions they catalyse keep on creeping into consciousness both as tools (don't forget to add the ATP to your ligations) and as the very objects of study; what does the recently discovered "homeoprotein" actually *do*? Yet we shy from the gory biochemical details. Why should this be so? Is it because of a lingering distaste for the blood and guts origin of the subject? Or because our algebra stopped short of the Michaelis-Menten equation, which we were led to think lay at the heart of the subject? Or because the chemical formulae were too big to remember, and the pathways too complex to comprehend except by rote the night before the exam? Or because explanations too often stopped before they really made sense? This is a topic that needs

exploring for almost all of molecular biology; this is common parlance among cell biologists. Yet molecular biology as I understand it has very little to do with the nucleus as such. It stands for a much more highly conceptualized scheme of things that transcends (for the time being) its cellular location. Actually, we don't even understand clearly why there is such a thing as a nucleus; bacteria get along without it and *Saccharomyces cerevisiae* commits mitosis without ever losing its nuclear membrane. Curiously, this kind of issue is not considered here. There is discussion of what the nuclear membrane lets in and out, of nuclear pores and how little we understand of how they work, of how the membrane vesiculates prior to mitosis (we have to wait for this to happen to get out!), but not of why it's there in the first place. (Alberts *et al.* in *Molecular Biology of the Cell* suggest that it's to keep protein synthesis out, thus allowing splicing to occur.) I was disappointed by de Duve's silence on the matter, for elsewhere he almost always has something interesting to say.

These two volumes are an extravagant buy for the individual, but they must go into my college library and I would advise intending biologists who have not done the subject at school to try and get hold of them before they go up to university. For, despite their deficiencies, the books are a splendid introduction to cell biology, and the first hundred pages could and probably should be read with profit and enjoyment by every biologist. □

Tim Hunt is a Lecturer in the Department of Biochemistry at the University of Cambridge.

Nature magnified

Peter Evennett

Single Lens: The Story of the Simple Microscope.

By Brian J. Ford.

Heinemann/Harper & Row: 1985.

Pp. 182. £10.95, \$14.95.

THE simple microscope, one whose optical system consists of a single lens, operating in the manner of a magnifying glass, has a distinguished place in history and quite remarkable capabilities, little known to the general scientist. The origins of its use are obscure, but by far the best known early work is that of Antony van Leeuwenhoek of Delft, communicated in the form of a series of letters to the Royal Society from 1673 until shortly before his death in 1723. One third of Brian Ford's book is devoted to Leeuwenhoek, his microscopes and specimens.

Using single-lens microscopes, Leeuwenhoek was the first to observe Protozoa, bacteria and spermatozoa: truly microscopic objects. In contrast Robert Hooke, Leeuwenhoek's contemporary working in London who published *Micrographia* in 1665, used a compound microscope (consisting of two lenses, objective and eyepiece). Both men's observations were excellent within their own limits, but while Hooke's work — on relatively large objects such as the flea or the sting of the nettle, for example — extended existing knowledge only slightly (though illustrating it most beautifully), Leeuwenhoek broke entirely new ground. The reason for this was principally his choice of microscope: lens design was then in its infancy and, to put it crudely, one poor lens gave better results than two poor lenses.

In addition to his letters, Leeuwenhoek also sent specimens, in small paper packets pasted to the letters; many of these survive in the archives of the Royal Society. In researching this book, Ford took some of the specimens to Utrecht and photographed them with an original microscope, demonstrating to those of us unlikely ever to lay hands on one the nature and quality of the image of a Leeuwenhoek microscope. A study of the surviving instruments shows that they were capable of resolving features as small as one micrometre, with a magnification of several hundred times — a performance as good as that of many school microscopes of today; indeed, for 300 years single-lens microscopes have been

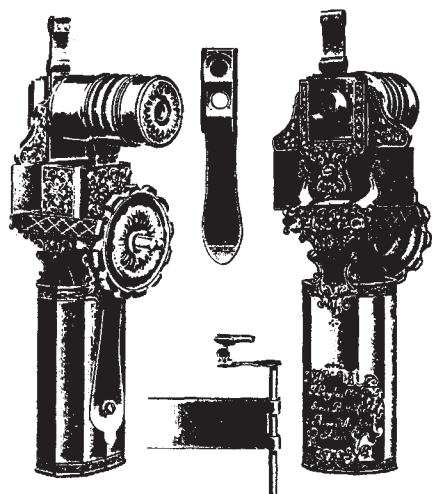
New in paperback

● The ninth edition of *Henderson's Dictionary of Biological Terms*, revised by Sandra Holmes. Publisher is Longman, price £4.50.

● *The Ecology of Animal Movement*, edited by Ian R. Swingland and Paul J. Greenwood. Publisher is Oxford University Press, price £12.50.

capable of revealing most of the significant structures that can be resolved by a modern light microscope.

After Hooke and Leeuwenhoek, the compound microscope increased in popularity as a "drawing-room" instrument, presumably because of its impressive appearance and greater ease of use; Ford describes the progress of the simple microscope in the form of the cheap flea-glass, the more versatile compass microscope, and the screw-barrel and later designs. A simple microscope is little more



Art of the instrument-maker — an elaborately decorated simple microscope produced in Paris, c. 1720.

than a device to hold a specimen and a lens in suitable relationship one to the other, its design naturally influenced by the nature of the specimen to be examined. Around the middle of the eighteenth century the "aquatic" microscope was designed, arranged to stand on a table, with a movable lens to enable it to scan sensitive creatures such as *Hydra* without disturbance. Linnaeus owned one of these, and a similar though superior example was used by the Scottish botanist Robert Brown in 1827 when he discovered "active molecules" (now known as Brownian movement), and a few years later when he described the nucleus and cytoplasmic streaming in the plant cell. Darwin took a microscope of this kind on the *Beagle* expedition in 1831, on Brown's advice, and wrote of his purchase of another in 1848. An essentially similar instrument is still available as a "dissecting microscope".

The simple microscope has not been well documented previously. Brian Ford's book is a clearly written non-technical account of the instruments, the lenses, their performance and use, illustrated with historical anecdotes, quotations and drawings, and with a useful bibliography. It will interest anyone who uses a microscope and many others besides. □

Peter Evennett is a Lecturer in the Department of Pure and Applied Zoology, University of Leeds, and Editor of the Proceedings of the Royal Microscopical Society.

Psychiatry: beyond analysis

Anthony W. Clare

The Broken Brain: The Biological Revolution in Psychiatry.

By Nancy C. Andreasen.

Harper & Row: 1985. Pp.278. \$16.95, £12.95.

A COMMONLY expressed dictum, usually by psychiatrists, is that there has been a revolution in psychiatry. Back in 1966, the late Sir Aubrey Lewis of the Maudsley Hospital wryly noted the common and invigorating habit of psychiatrists to rejoice "that they live in an age of rapid and impressive advance". Advances there have surely been but whether any of them warrants the description "revolutionary" is highly questionable.

Just after the Second World War, the new antipsychotic phenothiazines and the tricyclic antidepressants were hailed by psychiatrists such as William Sargant in Britain and Nathan Kline in the United States as the harbingers of a revolution in physical treatment which would empty the mental hospitals and re-establish psychiatry within the mainstream of general medicine. Even before the introduction of these drugs, however, the mental hospital population on both sides of the Atlantic had begun to fall and this trend too was acclaimed as reflecting a revolution in care. Much of its momentum derived from the realization that institutionalization of the chronic mentally ill could in many instances worsen the condition and add substantial secondary impairment. A vigorous new emphasis within psychiatry was born, social psychiatry, whose enthusiastic adherents emphasized the importance of energetic rehabilitation and rapid reintegration of the mentally ill within the general community. In the United States, the presidential involvement of John F. Kennedy gave the movement a particular momentum such that over a quarter of a century the mental hospital population has fallen by nearly 75 per cent (compared with the more pedestrian British decline of 25 per cent) and a whole chain of community mental health centres has been deployed across the continent.

The mid-1960s, a turbulent time within as well as outside psychiatry, witnessed the flowering and then the demise of arguably the most revolutionary idea of all, namely that mental illness does not actually exist but is a myth, a metaphor or merely an obfuscating and remarkably sticky label. The mentally ill were variously portrayed as outsiders, deviants and voyagers on a perilous journey of self-discovery and transcendentalism. Around the same time, the more prosaic yet more enduring advances in behavioural theory in turn yielded a series of promising techniques of be-

havioural treatment, including desensitization, operant conditioning and flooding, offering hope to sufferers of the most intractable of neurotic conditions such as phobic anxiety and obsessive compulsive disorder.

If one uses the term "revolution" in the sense of the undermining and overthrow of the established order, then none of these developments seriously threatened psychiatric orthodoxy. Psychotropic drugs, psychiatric rehabilitation, the development of community-based facilities, the application of behavioural techniques — each has been comfortably absorbed within the general mainstream of psychiatry. In the process, the original enthusiasm has been tempered by time and by bitter political experience. In the case of the much-vaunted community mental health centre programme, for example, professionals, politicians and the public appear deeply divided not about whether it was a good or a bad idea but about who precisely is responsible for the disaster that it has turned out to be. Its critics blame the haste, the ill-conceived design which has led to the warehousing of large numbers of chronically ill psychiatric patients in squalid and anti-therapeutic conditions within a community which clearly does not care, and the failure of its advocates to acknowledge the degree of disability manifested by those who are most seriously mentally ill. Its supporters attack the failure of Federal government to honour funding pledges made by its predecessors and the marked ideological shift in the way that Washington regards the concept of community welfare.

Unabashed, American psychiatrists are now acclaiming a new revolution. The author of *The Broken Brain*, a book revealingly subtitled *The Biological Revolution in Psychiatry*, is moved to declare with the enthusiasm of one who has witnessed the fall of some psychiatric Bastille, that "Psychiatry, like the prodigal son, has returned home to its place as a specialty within the field of medicine". For Nancy Andreasen, Professor of Psychiatry at Iowa College of Medicine, psychiatry is becoming increasingly scientific and biological in its orientation and medical science is now convinced that the serious forms of mental illness, such as schizophrenia and manic-depression, are due mainly to abnormalities in brain structure and chemistry.

British readers, familiar with the relatively close relationship between psychiatry and general medicine in this part of the world, may wonder what all the fuss is about and where the American prodigal son has been all these years. Yet the argument that in the United States at least there may indeed be a significant shift in psychiatric thinking, if not a revolution, does merit closer examination. There was a time, less than 25 years ago in fact, when it was difficult if not impossible to become the chairman of an American department of psychiatry unless one had undergone psychoanalytical train-

ing and a personal analysis. Up to quite recently, most major medical schools accepted psychoanalytically orientated psychotherapy as the foundation of psychiatric theory and practice, and in a number of them — most notably New York Medical College, Columbia, Pittsburgh, Tulane, Western Reserve and North Carolina — actual postgraduate psychoanalytical institutes were established. Not merely were most departmental heads psychoanalysts, so were most of the members of their departments, their psychiatric residents, their own personal psychotherapists, their wives and many of their government site visitors. In psychosomatic medicine, child and adolescent psychiatry, the psychotherapy of the neuroses, social casework, individual and group therapy, the impact and influence of psychoanalytic theory and practice were enormous.

Yet even at its zenith psychoanalysis was under fire. In 1959, two research psychiatrists, J. MacIver and F. Redlich, writing in the *Journal of the American Medical Association*, critically observed that "the hard truth is that we are expending enormous effort, time and money to train a highly specialized group of experts who will devote the bulk of their professional careers to a small group of highly selected patients". Several years later, a distinguished biologist, Seymour Kety, was appointed to head the psychiatric department at Johns Hopkins. Together with David Rosenthal and Paul Wender, he embarked on a programme of research which was to cast light on the chemical nature of synaptic transmission, develop a number of potent drugs capable of specific actions against many of the symptoms of severe mental illness, and provide new and substantial evidence of genetic factors operating in the aetiology of the major psychoses. Indeed one prominent psychoanalyst spoke prophetically at the time of Kety's appointment that it was calculated "to drive another nail into the coffin of analysis".

Despite these and related biological developments, however, American psychiatry clung on to its heavily analytical reputation until the past few years when suddenly it appears to be discarding it rapidly. Professor Andreasen expresses the widely held view that it is the exciting developments in neuroanatomy, neurochemistry, neuropathology, neuropharmacology, genetics and behavioural neurology which have not just undermined the pre-eminence of psychoanalysis but significantly eroded it. Her book, in a highly readable if occasionally simplistic fashion, emphasizes the manner in which such explanatory ideas as the catecholamine theory of depression and the dopamine theory of schizophrenia, while partial, flawed and even contradictory, have in their own way reminded a whole new generation of young psychiatrists of the possibilities and challenges inherent in the biological approach. Hemispheric localization, limbic system

function, GABA, positron-emission tomographic scanning, the dexamethasone suppression test, benzodiazepine receptors — the topics included and discussed, while in no way adding up to a coherent and plausible theory of brain dysfunction in psychiatric disease, illustrate the potential for research and the bewildering number of growing points.

In her enthusiasm for these developments, Andreasen, in common with many of the new biologists, ignores the lessons of recent history. The reverses suffered in turn by psychoanalysis and social psychiatry, each originally hailed as true revolutionary breakthroughs in the way we understand and treat mental illness, should not be ignored. If psychiatry in the United States has a fault it surely is its tendency to suffer from what Gore Vidal, in another capacity, has termed the peculiarly American affliction of "historical amnesia". At times, Professor Andreasen acknowledges the limitation of reducing all psychopathology to the level of abnormally firing neurones, all psychiatric disorder to the level of brain disease. On other occasions, her zeal gets the better of her and she commits precisely that error. Yet psychiatry has been here before, just over a century ago, but one will look in vain in the index for the names of Griesinger and Maudsley, Westphal and Meynert, men who also nourished the hope that one day severe psychopathology might be reduced to a matter of disturbed cerebral function.

The resurgence of biological enthusiasm doubtless owes much to the new and exciting plethora of scientific tools and techniques applicable to unravelling the complexity of brain function. It owes much too to the realization that psychoanalytic theory is sterile and psychoanalytic treatment is ineffective. But before participating in a single-minded and headlong rush to the brain, enthusiasts would do well to heed Kety's reminder that the language that we speak, the principles we value, the thoughts we think, the decisions that motivate our daily actions are to a greater extent a reflection of what we have learned than the result of how we are constituted. "The differences among us in respect of these qualities", he has recently written, "are not to be found in the structure or chemical composition of our brains but lie largely within the scope of the psychological and social sciences". If this is ignored, then crude biological reductionism is the result and the possibility that psychiatry might act as the bridge in medicine between the biological sciences on the one hand and the psychological and social sciences on the other hand evaporates. That is hardly a development worthy to be considered revolutionary. □

Anthony W. Clare is Professor in the Department of Psychological Medicine at St Bartholomew's Hospital Medical College, University of London.

Advertisement

The Giant Pandas of Wolong

GEORGE B. SCHALLER,
HU JINCHU, PAN WENSHI
& ZHU JING

Although the Chinese have known of the existence of the panda for several thousand years, the West became aware of its existence only in 1869, and it was not until 1978 that Chinese scientists first began to study the animal intensively in the wild. In 1980 China and the World Wildlife Fund began co-operative research on the species. *The Giant Pandas of Wolong* is not only the first report of this joint venture but the first detailed account of the panda's natural history. It describes the environment, distribution, life cycle and daily and seasonal activities of pandas in the wild and also helps to clarify the old scientific argument of whether the giant panda is related more closely to the bear or racoon family.

£26.25 Hardback 320pp with
32 halftones & 80 line drawings
0-226-73643-1 June

Neurohumoral Maintenance of Immune Homeostasis

ELENA A. KORNEVA,
VIKTOR M. KLIMENKO &
ELENORA K. SHKINEK
(translated by

Samuel A. Corson &
Elizabeth O'Leary Corson)

First published in Russian in 1978, this book provides evidence that the immune system is subject to influences from the central nervous system and the neuroendocrine system. The immune system is thus a critical link between mind and body in determining the onset and duration of many diseases, including cancer, infections and autoimmune processes. Korneva and her colleagues, pioneers in this field, here review a quarter century of research.

£33.50 Hardback 256pp
0-226-45042-2 June

Sonoran Desert Spring

JOHN ALCOCK

Sonoran Desert Spring is a celebration of the joys of observing desert plants and animals over the course of a season. A journal drawn from Alcock's studies of the Arizona desert, the book is natural history informed by modern evolutionary principles and guided by a genuine affection for a fragile land and its inhabitants.

£20.95 Hardback 196pp with
90 halftones & 6 colour plates
0-226-01258-1 July

CHICAGO

The University of Chicago Press
126 Buckingham Palace Road London SW1W 9SD

Sowing the seeds of awareness

David Streeter

Green Inheritance.

By Anthony Huxley.

Collins-Harvill/Anchor-Doubleday: 1985. Pp. 193. £9.95, \$19.95.

Our Green and Living World: The Wisdom to Save It.

By Edward S. Ayensu, Vernon H. Heywood, Grenville L. Lucas and Robert A. Defilippis.

Cambridge University Press/Smithsonian Institution Press: 1984. Pp. 255. £12.95, \$24.95.

In the Rain Forest.

By Catherine Caufield.

Heinemann/Knopf: 1985. Pp. 304. £10.95, \$16.96.

ONE of the most striking features of current radio and television, in Britain at least, is the number of natural history programmes that are finding their way into the schedules. Radio 4 alone, for example, is currently putting out five such programmes a week. Until recently, however, wildlife was considered to consist wholly of animals; even in David Attenborough's celebrated television series *Life on Earth*, plants made a significant appearance in only one out of the thirteen parts.

David Bellamy's early programmes have done much to redress the balance and coincided with the emergence of a higher profile conservation movement drawing attention to, among other issues, the attrition of the tropical rain forests. The established conservation bodies began to take notice. In 1980 the Fauna Preservation Society symbolically changed its name to the Fauna and Flora Preservation Society, and in March 1984 the World Wildlife Fund and the International Union for the Conservation of Nature and Natural Resources launched an international programme for plant conservation. As part of this campaign, two books, both aimed primarily at the general market, have appeared almost simultaneously. *Green Inheritance* is published as the official book of the World Wildlife Fund's Plants Campaign and is written by Anthony Huxley, a well-known writer and lecturer on plants. *Our Green and Living World* has an appropriately international authorship of an African, two Europeans and an American.

The purpose of both books is to arouse concern over the quite staggering rate of destruction of much of the Earth's natural vegetation and to point to the consequences. Inevitably one draws comparisons between the two. As a plea for plants the case put by Anthony Huxley is the more compelling. The factual material is well-researched, and while there are lapses here and there when the subject becomes more

physiological the text as a whole provides a fascinating and lucid account of mankind's relations with and dependence upon plants. Separate chapters deal with the role of plants as food, fuel, medicine and objects of beauty, as well as with the wider role of vegetation in atmospheric and soil processes. Accounts of the origins and spread of crop plants and the medicinal uses of plants are particularly good, the latter containing some quite astonishing information such as that on the uses or, more properly, abuses of the thornapple. Apart from Huxley's persuasive style, much of the success of *Green Inheritance* rests with its design. The combination of a continuous discursive text alongside short illustrated vignettes providing more detailed information works well, and the standard of graphics and art work is high.

Our Green and Living World covers much the same ground and overall, perhaps, is the more authoritative of the two. However, for the general reader it is likely that its message will be less compelling, the style sometimes being closer to that of an undergraduate text. Unlike *Green Inheritance* it includes eight chapters on different vegetation types from deserts to rain forests and from wetlands to oceans. It is also more comprehensive on the subjects of genetic conservation and under-exploited crops (the separate chap-

ters on firewood and cellulose are particularly good). Like *Green Inheritance*, the production can only be described as lavish, some of the photographs being quite superb.

The overall lesson of the two books is clear, and in particular both focus attention on the devastation of the tropical rain forests. Five years ago the US National Academy of Sciences calculated that over 50 million acres were destroyed or seriously damaged each year and the UN's Food and Agricultural Organization estimated in 1981 that at present rates almost one-fifth of the world's remaining rain forest would have gone by the end of the century. Catherine Caufield's *In the Rain Forest* is a carefully researched collection of essays on various aspects of rain forest ecology, exploitation and destruction. Subjects range from the building of the town and dam of Tukurui in the Amazon basin to the discovery of quinine. Figures and statistics spill off the pages, and the value of the book is enhanced by good indexing, text notes and a comprehensive bibliography, but it is not — and does not pretend to be — a textbook on rain forest ecology. For anyone unfamiliar with the rain forest this is undoubtedly a good read and an intentionally perturbing one. □

David Streeter is a Reader in the School of Biological Sciences, University of Sussex.

Who cares?

Marian Dawkins

Strategies of Being Female: Animal Patterns, Human Choices.

By Evelyn Shaw and Joan Darling.

Harvester/Walker: 1985. Pp. 164. £12.95, \$14.95.

IT WAS only in the last few pages of this book that I understood what the authors' purpose in writing it really was. Initially, I had thought it was going to be yet another polemic about the allegedly male-dominated sexist views of sociobiology in general and its treatment of sexual behaviour in particular. As I have never been able to understand this particular criticism of sociobiology (sociobiological emphasis on the evolutionary importance of female choice seeming, if anything, to over-emphasize the role of females), I began the book by being somewhat out of sympathy with the authors. Some distinctly group selectionist phraseology — such as "reproducing in anticipation of the species' future" (p.32) and "the true biological reasons for having children [are] to assure the survival of the human species" (p.121) — only increased my antipathy.

But I was curious. How were the rampaging, promiscuous, initiative-taking females they cite from all corners of the animal kingdom going to be related to our

own species, as promised in the introduction? Would promiscuous female fish or birds be used as a model for human beings? Would female hyenas, with their sham male genitalia, be held up as an example to us all?

Only in the last two chapters do Shaw and Darling argue that human females are in fact different from those of any other species. Females of other species may show a wide variety of sexual behaviour, but only human females can choose which particular one they will adopt for themselves. Looking after children is not, they argue, "biologically programmed in the genes of the female". "Words — not our hormones, and not some neurological Gestalt etched into our brains — teach us how to find a mate, who is appropriate as a mate and who isn't, when to have babies, how to give birth to them, and how to take care of them" (p.152). It is all the fault of culture, and the "myth of maternalism" in particular, that women are the ones most commonly adopting a child-caring role.

It is to attack this myth of maternalism — the claim that mothers are inherently better at caring for offspring than fathers — that turns out to be the book's main purpose. Men, it is argued, are just as good at looking after children as women are, if only they are given the chance. The trouble is they are so rarely given the chance. In a wide variety of cultures, women exclude men and refuse to allow them to be present at the birth of their own children. Shaw and

Darling argue that if only men could be present at the birth, they too would form strong bonds with their children and take on the care-taking role usually associated with mothers.

In support of this idea, they cite male wild-dogs, male marmosets and other species in which males take an active part in looking after the young. They also quote Rosenblatt's finding that male rats will

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Mother love — black-faced vervet monkey with infant.

begin to lick and retrieve young after they have been exposed to them for several days. The male's tendency to care for young is latent and revealed only if given the opportunity: "And just like a father rat, the human father is denied access to the environmental cues that would turn on his 'maternal' behaviour" (p.148).

There may or may not turn out to be something in all this. It is an interesting hypothesis, one which is certainly worth thinking about. What is disturbing to me is the way it is put over. Perhaps this is simply because the book is written in a popular style. Maybe words such as "babes" and "critters" have to be used to make sure readers do not feel overwhelmed by technical jargon; maybe a case has to be overstated in order to attract a public who would be bored by too much weighing up of evidence. But do we have to swing from male sexism to an equally unattractive female sexism? If we accept that women may wish and be able to fulfil a wider variety of roles than they often do, do we have to see society purely in terms of womens' choices and womens' destinies?

Shaw and Darling emphasize that "female" does not have to mean passive, meek or monogamous. But at the same time, perhaps unintentionally, they make female wishes and female roles dominant, without explaining why this kind of sexism is any more convincing — or any more acceptable — than the male kind. □

Marian Dawkins is Fellow in Biological Sciences at Somerville College, University of Oxford.

Ways to get back to the track

David Pimentel

State of the World 1985: A Worldwatch Institute Report on Progress Toward a Sustainable Society.

Project director Lester R. Brown.
W. W. Norton: 1985. Pp.301. Hbk \$22.75, £18.95; pbk \$9.50, £7.95.

IN PUBLISHING this detailed analysis of "the state of the world", the Worldwatch Institute has performed a valuable service for the public, government and industry. Examined in it are the interdependence of human population, global environment and world economy. Specifically, the authors assess the effect of population growth on deforestation, overgrazing, soil erosion, atmospheric pollution, and water and energy shortages. With this as a basis, they go on to consider the influence of natural resource degradation and resource shortages on the economy and standard of living of society.

Early in the book Brown and his colleagues tackle the question of whether climate change in some regions of the Earth may be population induced. They suggest that, indirectly, population growth may reduce rainfall because of overgrazing and other activities that remove or reduce vegetation cover. The increased exposure of bare soil increases the albedo: "Where this happens, as on the fringes of the Sahara, the affected areas reflect more heat into space". Canadian meteorologist, Kenneth Hare, who has investigated desertification in Africa, is quoted as concluding that "we seem to have arrived at a critical moment in the history of mankind's relation to climate. For the first time we may be on the threshold of man-induced climatic change". If indeed water shortages are population induced, then the decline can be reversed in some regions only by massive programmes in family planning, tree planting, soil and water conservation, and ecological-agricultural development on a huge scale.

The authors acknowledge the fact that major efforts to produce more food per hectare of land have slowed the growth of hunger, noting that agriculturalists have made use of biological diversity to breed new crop varieties that are more efficient in converting water and fertilizers into grain and other foods. At the same time, farmers have increased the amount of land cultivated and have greatly expanded the use of irrigation and fertilizers. Irrigation, for instance, has grown nearly three-fold since 1950 and fertilizer usage about nine-fold during much the same period. All these strategies have increased crop yields.

It remains to be seen, however, how far agricultural technology can go in producing more food from the world's limited

arable land resources. We already know that there are certain physical and chemical constraints on the capacity of plants to capture and fix solar energy into biomass. Recognizing such limitations, the authors question whether biotechnology will make the major changes necessary to continue improving yields. A major concern to emerge here is the effect that the rapid loss of biological diversity will have on the success of biotechnology and genetic engineering, but in placing emphasis on this point the authors neglect to explain that *most* species in nature are vital to the normal functioning of the ecosystem. For example, agriculture cannot function successfully with only its crop and livestock species; many organisms, especially the small organisms (such as microbes and insects) are also essential. Thus on a hectare of good land with ample water, the small organisms outweigh human biomass by at least 200 times and carry out the vital tasks of degrading wastes and chemical pollutants, and of recycling other resources.

In a chapter on energy efficiency, William Chandler reports on the important role that fossil energy plays in producing food, purifying water, protecting us from numerous kinds of diseases and transporting resources from one region to another. Several sound arguments support the view that conservation of energy and improved energy efficiency are necessary for economic improvement in both developing and developed countries. But although fossil energy is essential to our present standard of living, almost everyone agrees that new sources of energy are needed, preferably renewable in form. However, the chapter entitled "Harnessing Renewable Energy" is rather disappointing. Several of the sources cited lack scientific credibility. For example, one paper quoted here reports that producing ethanol from corn grain provides a net energy return. This is not so; in fact, the net energy loss is about 30 per cent greater than the ethanol energy produced.

Also highlighted is the idea of shifting from the corn-soybean crop mix now typical in the mid-west of the United States to a corn-sugar-beet mix. This proposed plan, which is supposed to supply all food needs for livestock plus up to 700 million barrels of ethanol, has numerous technological and agricultural flaws and has been

Conservation in Britain

In association with the publisher Webb & Bower, and sponsored by Gulf Oil, the Conservation Foundation have published *The Conservation Review*, their third annual report for "everyone concerned with preserving our heritage and natural environment".

The *Review* is intended to increase public awareness of environmental issues and draw attention to the activities of various organizations, local, national and international. Among the contributors are David Bellamy, Patrick Jenkin, Tony Soper and Jean Medawar. Price is hbk £10.95, pbk £4.95.

discounted by most energy specialists. Further, nowhere do the authors acknowledge the environmental risks associated with certain renewable systems. For example, the widespread burning of crop residues and dung as a fuel is severely intensifying soil erosion and water run-off, and is robbing the soil of essential nutrients. Because soil erosion and water run-off are a major feature of the *Report*, it is especially unfortunate that biomass energy systems are not identified as major causes of these problems. There are also drawbacks to hydropower, but again no mention is made of them. Dams and reservoirs have to be constructed to run hydropower systems, and in the United States, for instance, about 13,000 hectares of productive agricultural land has to be flooded per 1×10^9 kWh/yr of electricity produced (sufficient electricity for a city of 100,000). This is not to imply that hydropower is not a valuable energy system in the United States and many regions of the world, but that environmental risks should be evaluated to determine if the benefits for a region outweigh the costs.

One of the most challenging aspects of the resource and environmental problems facing us today is their "gradual, insidious nature". For example, a single rain storm may erode 1 mm of soil weighing 15 tonnes from a hectare. That loss may be imperceptible, but it does not take many rain storms and many years to remove the 2.5 cm of topsoil that generally requires 200 to 700 years to replace. As Brown and Wolf point out, "few countries will succeed in attempts to boost domestic food at the rate demanded by population growth if soil-depleting agriculture continues".

The authors of the *Report* clearly explain that soil, water, energy and biological resources are being threatened not only by the rapidly growing population but also by poor management practices in agriculture and other industries. Further, many excellent plans are suggested for tackling the problems. But, as they caution, long-term solutions will require the efforts of individual intellectuals and of politicians to get society back to the track — the intellectuals to identify the direction for change, and political leadership to devise and carry out the appropriate policies for stabilizing populations and effectively managing environmental resources.

State of the World is a thorough and up-to-date examination of the influence of mankind on the Earth's various resources. Particularly helpful and timely are the solutions to the complex problems that are identified in the book. This is an exceptionally well-written report which scientists and industrial and political leaders would do well to read and to act upon. □

David Pimentel is Professor of Insect Ecology and Agricultural Science at the College of Agriculture and Life Sciences, Cornell University.

Biology as science

A.J. Cain

The Cambridge Encyclopedia of Life Sciences.

General editors Adrian Friday and David S. Ingram.

Cambridge University Press: 1985.

Pp.432. £25, \$45. To be published on 9 May.

TO TRY to encompass within a volume of 432 pages a complete encyclopaedia of the life sciences strikes one as a most noble aim and a forlorn hope. The editors and contributors have succeeded in this book to a really remarkable degree.

The volume is divided into three parts: "Processes and Organisation" (the cell, the organism, genetics, behaviour and sociobiology, ecology); "Environments" (marine, coastal, terrestrial, freshwater and wetland, living organisms as environments); and "Evolution and the Fossil Record" (the evolutionary process, palaeontology, early events in evolution, origin and development of the land flora and fauna, recent history of the flora and fauna). A classification of living organisms, a species index (generic also) and a subject index complete the book.

The contents page refers under Part 2 to what appears to be an unnumbered chapter on biogeographical zones. The relevant pages in fact carry a set of maps only, and neither biogeography nor any names of biogeographical regions appear in the subject index. There are, however, excellent accounts of plate tectonics, numerous maps of the drifting continents, and much information on regional history and animal diversity in the section on the origin and development of the land flora and fauna, the best place to discuss the development of biogeographical diversity.

The book is written in a very readable style and has obviously been most carefully corrected in proof — one of the very few books I have seen in recent years that can be put into the hands of the young without a warning against bad spelling and gram-

mar. It is a model of exposition, excellently illustrated. In the words of the editors:

The Cambridge Encyclopedia of Life Sciences is not a natural history cataloguing the living world, it does not attempt to argue a case for or against biological conservation, nor does it consider the yet greater ethical problems which arise from the genetic manipulation of cells. Instead, the *Encyclopedia* seeks to present biology as a science, and in addition constitutes a plea for considering all biological phenomena as indissolubly linked, a circumstance which underlies both the frustration and the fascination felt by all life scientists.

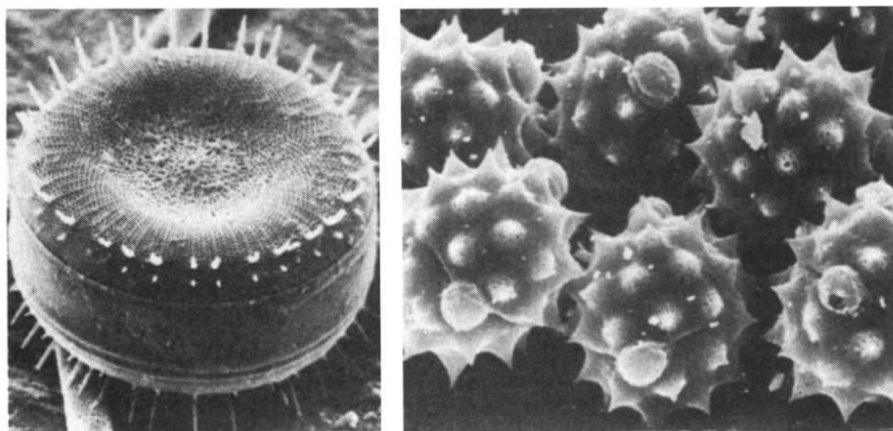
Inevitably, the necessary compression means that many topics can only be treated summarily, as a series of key concepts, their interrelations, and a brief mention of their empirical basis. In some ways this is a positive advantage, forcing the writers to get to the real bones of the subject. One could give a series of lectures on the single paragraph (p.296):

Some biologists have argued that if we can reconstruct the pattern of evolution, the *phylogeny*, this will provide insight into the ways in which evolutionary change occurs. The opposite point of view is that knowledge of the processes of evolutionary change is essential before it is possible to find out the pattern of relationships between organisms.

Yes, indeed!

Equally inevitably, specialists will be dissatisfied because so much can only be mentioned. It would take a team of reviewers as large as the team of authors and editors (31, 22 of them in Cambridge) to deal with the chapters fully. In my own speciality, I notice only the total absence of ring species and allopolyploid plants from the section entitled "Mechanisms of Speciation", in which too much space is given to the latest fads in speciation theory. As against that, what a pleasure it is to see a proper discussion of parasites in the section on living organisms as environments. Surely there are more species of parasitic than of free-living organisms, yet this major aspect of evolution and the ecosystems of the world is usually mentioned merely in passing except in highly specialized texts. □

A.J. Cain is Derby Professor of Zoology at the University of Liverpool.



Revelations of the scanning electron microscope — the diatom *Stephanodiscus* (left) and pollen of *Chrysanthemum mycoais*.

From The Cambridge Encyclopedia of Life Sciences

A matter of taste

Nicholas Kurti

On Food and Cooking: The Science and Lore of the Kitchen.

By Harold McGee.

Charles Scribner's Sons: 1984. Pp.684. \$29.95.

How refreshing it is to be treated to an inspiring blend of botany, zoology, physiology, anatomy, biochemistry, and, of course, gastronomy. In his introduction to what is almost a scientific, literary and historical encyclopaedia of food and its enjoyment, Dr McGee admits to the importance of experience, intuition and a discriminating palate in preparing food, but he also asserts that "a general understanding of what is going on in the pan ... can compensate for a lack of familiarity with particular ingredients and techniques". The book is studded with examples of the beneficial intrusion of science into the kitchen, often expressed in vivid terms, such as the remark that "roast and meringue are different outcomes of the same process [but] the knowledge of protein coagulation need not diminish the pleasure in consuming a well-made roast or a meringue". McGee also reminds us that "while we may be up on Black Holes and the Big Bang and the

Primordial Soup we can't explain how flour thickens soup".

Nearly four-fifths of the book is devoted to accounts of the foodstuffs that go into our diet, grouped according to their origin rather than their chemistry. Scientific explanations are interspersed with historical remarks and tips on how to avoid or rectify cooking disasters and there is no lack of debunking (or justifying, as the case may be) of culinary and dietary fads. For example the much-vaunted blessings of fibre and bran are discussed critically but with tolerance. The author recognizes the nutritional value of whole-grain bread, but also quotes the rickets epidemic in Dublin during the Second World War when rationing of dairy products led to a low calcium intake, the assimilation of which was inhibited by the high bran content of the diet.

In discussing the progress of meat from the slaughterhouse to the table, McGee emphasizes the need to keep the animal free of stress before slaughter; muscular tension depletes the glycogen supply, less lactic acid accumulates in the muscle after slaughter and the result is inferior "dark-cutting" meat. The use of tenderizers based on naturally occurring proteolytic enzymes is also described but, surprisingly, the use of the hypodermic syringe to obtain even distribution of the tenderizer is not mentioned. And although it is true that these

enzymes work faster at higher temperatures, I have found that fresh pineapple juice will tenderize pork chops at room temperature in 2-4 hours.

There is too an excellent explanation of the conflicting demands on the *rôtisseur* trying to achieve the conversion of the collagen in connective tissues into gelatin (a slow process) without drying out and toughening the muscle fibres. There is, however, nothing about the low-temperature (less than 75°C) roasting technique pioneered by Sir Benjamin Thompson, Count Rumford (1753-1814), the Anglo-American soldier, statesman and natural philosopher who was also a culinary innovator. This method works well: I proved its excellence to my own satisfaction — and, apparently, to the satisfaction of those

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

who ate the joint — by roasting a 4lb leg of lamb for 8 hours in a slow oven.

McGee gives a full account of the arguments about the use of copper bowls when making meringues, culminating in a re-run of the beautiful piece of work undertaken at Stanford University by McGee and his colleagues (*Nature* 308, 667; 1984) which showed conclusively that traces of copper ions do indeed make egg-white foam more durable through the formation of a copper-protein complex. It is a pity that neither this book nor the original article indicates the minimum amount of copper needed per egg. It is bound to be very small and phials of "meringue-improvers" could be sold at a price many thousand times the value of their contents. (I found it odd, incidentally, that albumen coagulation (which occurs at around 60°C) is regarded as an essential feature of meringues; it is well known that meringues can be made at considerably lower temperatures, even at 20°C, especially if reduced pressure is used.) There are many other useful tips for the amateur confectioner; here at last, for instance, we have a clear explanation of how peppermint cream gets into the thin crisp chocolate of after-dinner mints.

It is in the chapter on sauces that the enthusiasm of the gastronome, the artistic skill of the cook and the critical approach

of the scientist shine most brightly. This is a veritable *saucier's* vade-mecum. All the traditional tricks are explained, new techniques are suggested and there is also, among other good things, a full account of the saga of the Resurrection of a Curdled Sauce Béarnaise which began in 1977 with an article in *Nature* (270, 572) by C.M. Perram, C. Nicolau and J.W. Perram.

More prosaically, in writing about food additives McGee reminds those with a taste for nostalgia of "bread mixed up with chalk, alum and bone ashes", of "greens boiled with brass half-pence" and "butter manufactured with candlegrease and kitchen stuff", practices which were prevalent in mid and late eighteenth century England. Smollet bemoaned the fact that "these enormities might be remedied, but the wise

patriots of London have taken into their heads that all regulation is inconsistent with liberty". We are better off now, inasmuch as food adulteration on the scale of the past can no longer be perpetrated with impunity, and the author illustrates, with the help of the cases of saccharin and sodium nitrite, the difficulties in arriving, on the basis of often incomplete scientific evidence, at solutions which satisfy public demand and reduce the risk to public health. Similar good sense characterizes the two brief sections "Food and the Body" and "The Principles of Cooking", though I was surprised that nowhere in the book are calories

mentioned. I should liked to have seen at least a modest table giving the values of the principal foods. This omission may have been deliberate, however, the author wanting to emphasize in this way a view, latent throughout the text and expressed felicitously at the end of the chapter on nutrition, that "Nutrition by the numbers is no guarantee of the good life and probably precludes it by overriding the calls of appetite and pleasure".

Dr McGee's book brings to mind a remark of Count Rumford:

The advantage that would result from an application of the late brilliant discoveries in philosophical chemistry, and other branches of natural philosophy of mechanics, to the improvement of the art of cookery, are so evident and so very important that I cannot help flattering myself that we shall soon see some enlightened and liberal-minded person of the profession take up the matter in earnest, and give it a thoroughly scientific investigation.

In what art or science could improvements be made that would more powerfully contribute to increase the comforts and enjoyments of mankind.

Two centuries later, with the publication of *On Food and Cooking*, gastronomes and scientists alike can rejoice that Rumford's hopes are on their way to fulfilment. □

Nicholas Kurti is Emeritus Professor of Physics at the University of Oxford.

Mathematics

Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics. By DAVID HESTENES and GARRET SOBCZYK. Reidel: 1984. Pp.314. ISBN 90-277-1673-0. Np.

Incline Algebra and Applications. By Z-Q CAO, K.H. KIM and F.W. ROUSH. Ellis Horwood: 1984. Pp.165. ISBN 0-85312-765-4. £16.50.

Mathematical Aspects of Superspace. H.J. SEIFERT, C.J.S. CLARKE and A. ROSENBLUM (eds). Reidel: 1984. Pp.214. ISBN 90-277-1805-9. Np.

W.G. Cochran's Impact on Statistics. PODURI S.R.S. RAO and JOSEPH SEDRANSK (eds). Wiley: 1984. Pp.431. ISBN 0-471-09912-0. £46.50.

Astronomy

Armchair Astronomy. By PATRICK MOORE. Patrick Stephens, Denington Estate, Wellingborough, Northants NN8 2QD. Pp.185. ISBN 0-85059-718-8. £9.95.

Astronomical Objects for Southern Telescopes: A Handbook for Amateur Observers. By E.J. HARTUNG. Cambridge University Press: 1984. Pp.237. Pbk ISBN 0-521-31887-4. Pbk £8.95.

Astronomy and Astrophysics Abstracts Vol.35/36. S. BOHME, W. FRICKE, H. HEFELE, I. HEINRICH, W. HOFMANN, D. KRAHN, V.R. MATAS, L.D. SCHMADEL and G. ZECH (eds). Author and Subject Indexes to Vols 25-34 Literature 1979-1983 prepared by I. HEINRICH and L.D. SCHMADEL. Springer-Verlag: 1984. Pp.891. ISBN 3-540-13651-7. DM 189, \$68.80.

Atlas of the Night Sky. STORM DUNLOP (ed.). Newnes Books: 1984. Pp.80. ISBN 0-600-35113-0. £4.95.

Clusters and Groups of Galaxies. F. MAR-DIROSSIAN, G. GIURICIN and M. MEZZETTI (eds). Reidel: 1984. Pp.659. ISBN 90-277-1772-9. Np.

Collins Guide to Stars and Planets. By IAN RIDPATH. Collins: 1984. Pp.384. ISBN 0-00-219071-0. £8.95.

Comets: A Descriptive Catalog. By GARY W. KRONK. Enslow: 1984. Pp.331. Pbk ISBN 0-89490-072-4. Pbk £15.50.

Mathematical Astronomy in Copernicus's de Revolutionibus. Parts 1 & 2. By N.M. SWERDLOW and O. NEUGEBAUER. Springer-Verlag: 1984. Part 1 pp.537, ISBN 0-387-90939-7/3-540-90939-7. Part 2 pp.711, ISBN 0-387-90939-7/3-540-90939-7. DM 238, \$88.80.

Relativistic Astrophysics and Cosmology. X. FUSTERO and E. VERDAGUER (eds). Wiley: 1984. Pp.301. ISBN 9971-966-60-3. £32.25.

Serendipitous Discoveries in Radio Astronomy. K. KELLERMANN and B. SHEETS (eds). National Radio Astronomy Observatory, PO Box 2, Green Bank, WV 24944: 1984. Pp.321. Np.

Physics

Advances in Applied Mechanics. JOHN W. HUTCHINSON and THEODORE Y. WU (eds). Academic: 1984. Pp.376. ISBN 0-12-002024-6. \$89, £67.

Applied Atomic Collision Physics. Vol. 2 Plasmas. C.F. BARNETT and M.F.A. HARRISON (eds). Academic: 1984. Pp.500. ISBN 0-12-478802-5. \$80, £62.

Basic Plasma Physics II. Handbook of Plasma Physics, Vol. 2. A.A. GALEEV and R.N. SUDAN (eds). North-Holland: 1984. Pp.850. ISBN 0-444-86645-0. Dfl 495, \$190.50.

Computational Methods of Neutron Transport. By E.E. LEWIS and W.F. MILLER. Wiley: 1984. Pp.401. ISBN 0-471-09245-2. £49.80.

Elementary Particles. By RAMON PRASAD. Geometrica Press, 21 Nassington Road, London NW3 2TX, UK: 1984. Pp.76. Pbk ISBN 0-904377-02-4. Pbk £5.

Equilibrium Properties of Fluids and Fluid Mixtures. By ALEKSANDER KREGLAWSKI. Texas A&M University Press: 1984. Pp.253. ISBN 0-89096-183-2. \$32.50.

General Relativity. By ROBERT M. WALD. University of Chicago Press: 1984. Pp.491. Pbk ISBN 0-226-87033-2. Pbk £25.50, \$30.

The History of Classical Physics. By R.W. HOME. Garland: 1984. Pp.324. ISBN 0-8240-90607-5. \$53.00.

Mathematical Physics Reviews. Vol. 4. S.P. NOVIKOV (ed.). Harwood: 1984. Pp.280. ISBN 3-7186-0146-X. \$99.50.

Modern Methods of Particle Size Analysis. HOWARD G. BARTH (ed.). Wiley: 1984. Pp.309. ISBN 0-471-87571-6. £58.50.

Molecular Quantum Electrodynamics: An Introduction to Radiation-Molecule Interactions. By D.P. CRAIG and T. THIRUNAMACHANDRAN. Academic: 1984. Pp.324. ISBN 0-12-195080-8. Np.

Nuclear Tracks and Radiation Measurements: Solid State Nuclear Track Detectors. G. ESPINOSA *et al.* (eds). Pergamon: 1984. Pp.651. ISBN 0-08-031420-1. £68.75, \$110.

Biological Sciences

Ionic Channels of Excitable Membranes. By BERTIL HILLE. Sinaur: 1984. Pp.426. ISBN 0-87893-322-0. \$32.50.

Laboratory Guide to Insect Pathogens and Parasites. By GEORGE O. POINAR and GERARD M. THOMAS. Plenum: 1984. Pp.392. ISBN 0-306-41680-8. \$49.50.

Lecture Notes on Coastal and Estuarine Studies. Vol. 8 Marine Phytoplankton and Productivity. O. HOLM-HANSEN, L. BOLIS and R. GILLES (eds). Springer-Verlag: 1984. Pp.175. Pbk ISBN 3-540-13333-X/0-387-13333-X. Pbk DM 36, \$14.20.

The Lesser Apes: Evolutionary and Behavioural Biology. HOLGER PREVSCHOFT *et al.* (eds). Edinburgh University Press: 1984. Pp.709. ISBN 0-85224-448-7. £65.

Leukemia. E. THIEL and S. THIERFELDER. Springer-Verlag: 1984. Pp.305. ISBN 3-540-13289-9/0-387-13289-9. Np.

Lipid Research Methodology. JONA. STORY (ed.). Alan R. Liss: 1984. Pp.296. ISBN 0-8451-1659-2. £37.

Lipids in Plants and Microbes. By J.L. HARWOOD and N.J. RUSSELL. George Allen & Unwin: 1984. Pp.162. Hbk ISBN 0-04-574021-6; pbk ISBN 0-04-574022-4. Hbk £19.50; pbk £7.95, \$12.95.

Membrane Dynamics and Transport of Normal and Tumor Cells. L. TRÖN *et al.* (eds). Akadémiai Kiadó, Budapest: 1984. Pp.387. ISBN 963-05-3889-X. £22.75.

Membranes, Channels, and Noise. ROBERT S. EISENBERG, MARTIN FRANK and CHARLES F. STEVENS (eds). Plenum: 1984. Pp.288. ISBN 0-306-41806-1. \$47.50.

Methods in Enzymology. Vol. 105 Oxygen Radicals in Biological Systems. LESTER PACKER (ed.). Academic: 1984. Pp.600. ISBN 0-12-182005-X. £67.50, £52.

Methods in Enzymology. Vol. 107 Posttranslational Modifications, part B. FINN WOLD and KIVIE MOLDAVE (eds). Academic: 1984. Pp.688. ISBN 0-12-182007-6. \$69.50, £49.

Methods in Marine Zooplankton Ecology. By MAKOTO OMORI and TSUTOMU IKEDA. Wiley: 1984. Pp.332. ISBN 0-471-80107-0. £47.80.

Methods in Microbiology. Vol. 14. T. BERGAN (ed.). Academic: 1984. Pp.228. ISBN 0-12-521514-2. \$40, £27.

Methods in Pharmacology. Vol. 5 Myocardial Biology. ARNOLD SCHWARTZ (ed.). Plenum: 1984. Pp.236. ISBN 0-306-41684-0. \$42.50.

Mode of Action of Antifungal Agents. A.P.J. TRINCI and J.F. RYLEY (eds). Cambridge University Press: 1984. Pp.405. ISBN 0-521-26171-6. £42.50.

The Molecular Biology of Adenoviruses 3. Current Topics in Microbiology and Immunology. WALTER DOERFLER (ed.). Springer-Verlag: 1984. Pp.110. ISBN 3-540-13138-8. DM 62, \$24.30.

The Molecular Virology and Epidemiology of Influenza. CHARLES H. STUART-HARRIS and C.W. POTTER (eds). Academic: 1984. Pp.305. ISBN 0-12-674740-7. \$29.50, £19.50.

Monoamine Innervation of Cerebral Cortex. Vol. 10 Neurology and Neurobiology. LAURENT DESCARIES, TOMAS R. READER and HERBERT H. JASPER (eds). Alan R. Liss: 1984. Pp.361. ISBN 0-8451-2711-X. £52.

Motility and Taxis in Prokaryotes. By A.N. GLAGOLEV. Harwood: 1984. Pp.279. ISBN 3-7186-0160-5. \$126.

Movement Disorders: Tremor. LESLIE J. FINDLEY and RUDY CAPILDEO (eds). Macmillan, London: 1984. Pp.493. ISBN 0-333-35040-5. £60.

Neartic Avian Migrants in the Neotropics. By JOHN H. RAPPOLE *et al.* World Wildlife Fund: 1983. Pp.646. Np.

Neonatal Behavioral Assessment Scale. By T. BERRY BRAZELTON. Blackwell Scientific: 1984. Pp.125. ISBN 0-632-01263-3. £7.50.

Neural Mechanisms of Startle Behavior. ROBERT C. EATON (ed.). Plenum: 1984. Pp.377. ISBN 0-306-41556-9. Np.

The Neurobiology of Pain. ARUN V. HOLDEN and WILLIAM WINLOW (eds). Manchester University Press: 1984. Pp.414. ISBN 0-7190-1061-6. £35.

Neurotransmitter Receptors: Mechanisms of Action and Regulation. SHOZO KITO *et al.* (eds). Plenum: 1984. Pp.297. ISBN 0-306-41729-4. Np.

Neurotransmitters and the Cerebral Circulation. E.T. MACKENZIE, J. SEYLAZ and A. BÈS (eds). Raven: 1984. Pp.270. ISBN 0-88167-010-3. \$54.50.

Nonlinear Electrodynamics in Biological Systems. W. ROSS ADEY and ALBERT F. LAWRENCE (eds). Plenum: 1984. Pp.603. ISBN 0-306-41736-7. \$89.50.

Organometallic Compounds and Living Organisms. By JOHN THAYER. Academic: 1984. Pp.273. 0-12-686060-7. \$49.50, £38.50.

General

Sea Bed Mechanics. By J.F.A. SLEATH. Wiley: 1984. Pp.335. ISBN 0-471-89091-X. £46.

The Second Law. By P.W. ATKINS. W.H. Freeman: 1984. Pp.230. ISBN 0-7167-5004-X. £16.50.

Sheets of Many Colours. By GORDON L. HERRIES DAVIES. Royal Dublin Society: 1983. Pp.242. ISBN 0-86027-014-9. IR£15.

Sintering and Heterogeneous Catalysis. Materials Science Research, Vol. 16. G.C. KUCZYNSKI, ALBERT E. MILLER and GORDON A. SARGENT (eds). Plenum: 1984. Pp.349. ISBN 0-306-41666-2. Np.

Slope Instability. DENYS BRUNSDEN and DAVID B. PRIOR (eds). Wiley: 1984. Pp. 620. ISBN 0-471-90348-5. £21.50.

The Smith Conjecture. JOHN W. MORGAN and HYMAN BASS (eds). Academic: 1984. Pp.243. ISBN 0-12-506980-4. \$49.50, £35.

Source Book of Geriatric Assessment. Vol. 1, Evaluations in Gerontology; Vol. 2, Review of Analysed Instruments. By LILIANE ISRAËL, DJORDJE KOZAREVIĆ and NORMAN SARTORIUS. Karger: 1984. Vol. 1: pp.439, ISBN 3-8055-3830-8; Vol. 2: pp.226, ISBN 3-8055-3831-6. SwFr. 374, \$224, DM 448.

Sparse Matrix Technology. By SERGIO PISSANETSKY. Academic: 1984. Pp.321. ISBN 0-12-557580-7. \$55, £32.50.

Special Issue on Historical Whaling Records. M.F. TILLMAN and G.P. DONOVAN (eds). International Whaling Commission: 1984. Pp.269. ISBN 0-906975-11-5. Np.

The Structure of Television. By P. GOULD, J. JOHNSON and G. CHAPMAN. Pion: 1984. ISBN 0-85086-110-1. £18.

The Structure of Time. By W.H. NEWTON-SMITH. Routledge & Kegan Paul: 1984. Pp.262. Pbk ISBN 0-7102-0389-6. Pbk £7.95.

Systems Analysis and Modelling in Defense: Development, Trends and Issues. REINER K. HUBER (ed.). Plenum: 1984. Pp.913. ISBN 0-306-41609-3. \$125.

Systems: Concepts, Methodologies, and Applications. By BRIAN WILSON. Wiley: 1984. Pp.339. ISBN 0-471-90443-0. £15.95.

Testament to the Bushmen. By LAURENS VAN DER POST and JANE TAYLOR. Viking: 1984. Pp.176. ISBN 0-670-80065-1. £12.95.

Theorie der Lebenserscheinungen. By HERMANN SCHALTEGGER. S. Hirzel Verlag, Stuttgart: 1984. Pp.306. ISBN 3-7776-0395-3. DM 98.

Timing and Time Perception. JOHN GIBBON and LORRAINE ALLAN (eds). New York Academy of Sciences: 1984. Pp.654. Hbk ISBN 0-89766-240-7; pbk ISBN 0-89766-241-5. Pbk \$144.

Dynamic solidification of a binary melt

Herbert E. Huppert* & M. Grae Worster†

* Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK

† Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The fundamental physics and fluid dynamics of a solidifying two-component system cooled from below is analysed quantitatively. Theoretical results for the rate of growth of the solid are in good agreement with experimental measurements on crystallizing aqueous solutions. The results are applicable to the solidification of a wide variety of binary alloys.

THE cooling and resultant solidification of a melt is important in many different fields, including crystal growth^{1,2}, electrical engineering^{3,4}, geology^{5,6}, geophysics^{7,8}, metallurgy^{9,10} and oceanography^{11,12}. The investigation of the fluid dynamic processes which often dominate solidification has been advanced by the suggestion¹³⁻¹⁵ that the mechanisms may be elucidated by laboratory experiments on the freezing of aqueous solutions. In some cases^{14,15} the studies were aided greatly by the concurrent analysis of a predictive mathematical model. Here, we use this powerful approach and describe some simple experimental and associated theoretical models for the cooling and crystallizing at a horizontal boundary of an initially homogenous fluid. The aim is to highlight the fundamental phenomena and to minimize consideration of inessential details by studying a simple and straightforward situation. We use aqueous solutions because they are easy to handle in the laboratory, but the concepts developed are applicable directly to the solidification of a wide variety of binary alloys.

First, we discuss briefly the phase diagram of the solid/melt mixture and then concentrate on effects associated with cooling and solidifying a melt from below. We determine theoretically the rate of solidification if the solid/melt interface is perfectly flat and horizontal, and also evaluate the criterion for which the interface becomes unstable. In a series of laboratory experiments, we have grown ice with unstable interfaces from different aqueous solutions and we have developed a simple theoretical model whose predictions are in good agreement with the observations. Our main quantitative conclusions are theoretical and experimental relationships for the rate of growth of the crystal block, the total volume of solid produced and an expression for the volume fraction of the solid product.

Phase diagram

The chemical composition of the phase obtained by the solidification of a melt with two chemical components is determined at thermodynamic equilibrium by the phase diagram (Fig. 1). This diagram represents the chemical compositions of melt and solid in equilibrium with each other by the liquidus and solidus respectively. At a temperature and composition represented by a point above the liquidus, the system is liquid. Between the liquidus and the eutectic line, the solidus defines the composition of the material solidifying from the melt, and in general specifies solid phases whose composition differs from both the melt and the pure components of the system^{6,9}. For most aqueous solutions and many binary alloys, however, the phase diagram has the special form illustrated in Fig. 1a, where solidification from a melt (solution) with sub-eutectic composition yields a solid phase consisting of a pure component (ice). It is usually the case that density in the melt is a much stronger function of composition than of temperature (see Fig. 1), which plays an important role in the fluid dynamics of solidification.

The solidification of a melt of eutectic composition, yielding solid of the same bulk composition, is equivalent to the solidification of a pure material. Otherwise, study of the process of solidification requires that melts with compositions greater and less than that of the eutectic must be dealt with separately, while

the direction of cooling (from above or below) also requires distinctive treatment. Table 1 shows the 3×2 matrix by which the problems of solidification may be classified. What follows concentrates mainly on cooling from below of a melt whose composition is below the eutectic value.

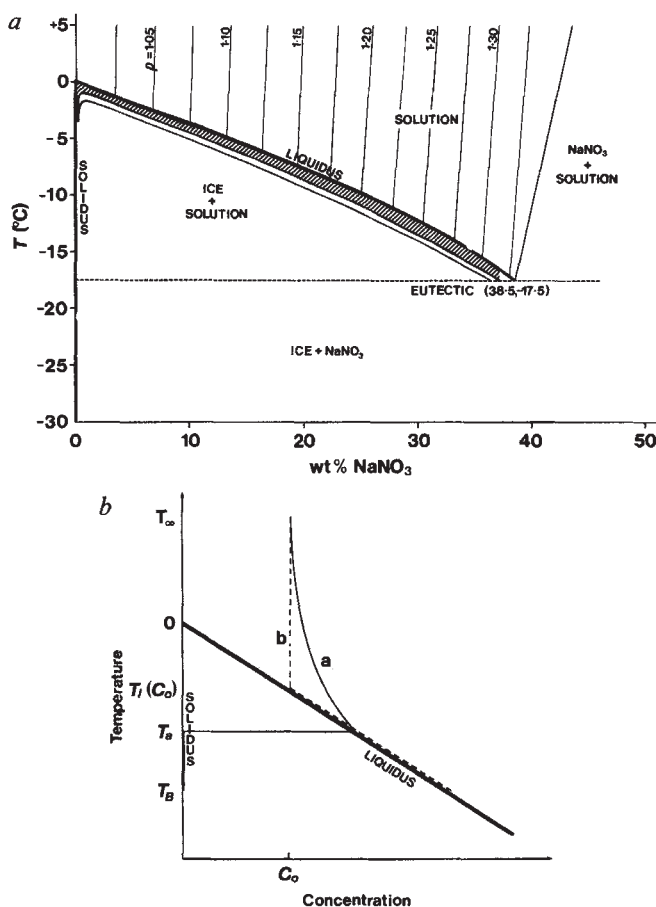


Fig. 1 a, Phase diagram for an aqueous solution of NaNO₃ (refs 20, 24) with axes of temperature and concentration. The same diagram acts as a schematic for a general binary alloy with components A and B. The figure also presents the lines of constant density with the density values attached. Hatched region, morphologically stable region for $T_\infty = 15$ °C. The curve just beneath the hatched region indicates the curve of marginal stability for $T_\infty = 30$ °C. In determining the curves the following values of the constant physical parameters have been used: $C_{pm} = 1.0$ cal g⁻¹ °C⁻¹, $C_{ps} = 0.48$ cal g⁻¹ °C⁻¹, $D = 10^{-5}$ cm² s⁻¹, $L = 80$ cal g⁻¹, $\kappa_m = 0.0013$ cm² s⁻¹, $\kappa_s = 0.012$ cm² s⁻¹, $\rho_s = 0.916$ g cm⁻³. Further, the values of Γ and ρ_m were obtained as a function of C_0 from ref. 20. b, Phase diagram used here, with temperature and concentration axes chosen so that a solid of zero concentration forms at zero temperature. Heavy solid line, straight-line approximation to the liquidus. Thin curve a, temperature against concentration profile in the morphologically stable regime. Dashed curve b, approximate profile used in our model of the morphologically unstable regime.

Table 1 The six different cases that arise when an initially homogeneous solution is cooled at a horizontal boundary

	Cool from		
	Above	Below	
$C > C_E$	13	25	Light fluid released
$C = C_E$	13	14	Uniform composition throughout
$C < C_E$	22	here	Heavy fluid released
	Thermally unstable	Thermally stable	

Numbers refer to the reference in which each case is discussed. See text for explanation of symbols.

Eutectic growth

The solidification of a melt of eutectic composition involves only variations in temperature, because the composition in the melt and solid is constrained to remain at the eutectic value. This situation is equivalent to one described by Carslaw and Jaeger¹⁶, who present the analytical solution for a column of melt of infinite height whose initial temperature is T_∞ and which is cooled at a lower boundary whose temperature is T_B . We conducted two experiments using eutectic compositions of aqueous Na_2CO_3 (eutectic composition $C_E = 5.7$ wt%, eutectic temperature $T_E = -2.1^\circ\text{C}$) in an insulated Perspex tank $20\text{ cm} \times 20\text{ cm} \times 45\text{ cm}$ high. Our observations of the height of the compact eutectic solid as a function of time in one of our experiments are plotted in Fig. 2. They are in good agreement with the analytical solution, which attests to the reliability of our experimental methods.

Sub-eutectic growth

We have extended the study of the solidification of a eutectic melt to the case where the composition is not that of the eutectic but in which the initial uniform concentration of solute C_0 is less than C_E . Solid of composition $C = 0$ grows on the cooled boundary and, to begin with, we assume that the solid/melt interface remains flat and parallel to the boundary. The fluid released by the crystallization is relatively heavy and thus remains just above the interface, so that the concentration gradient which develops is stable. Thus, no physical motion takes place and the transport of heat and solute is by molecular diffusion alone. The governing equations are then those for heat transport (in one dimension, taken as z measured upwards from the cooled boundary) above and below the interface and for solute diffusion in the melt. The equations for the temperature $T(z, t)$ and the composition $C(z, t)$ are

$$\rho_s C_p \frac{\partial T}{\partial t} = k_s \frac{\partial^2 T}{\partial z^2} \quad (\text{for } z < h(t), \text{ the position of the moving boundary}) \quad (1)$$

$$\rho_m C_p \frac{\partial T}{\partial t} = k_m \frac{\partial^2 T}{\partial z^2} \quad (2)$$

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial z^2} \quad (z > h(t); \text{ that is, for } z \text{ in the melt}) \quad (3)$$

$$T = T_B \quad (z = 0) \quad (4)$$

$$T \rightarrow T_\infty, \quad C \rightarrow C_0 \quad (z \rightarrow \infty \text{ or } t \rightarrow 0) \quad (5)$$

The physical parameters are the solute diffusivity D , the thermal conductivity k , the density ρ and the specific heat C_p , all of which are considered to be constant. Subscripts s and m refer to properties of the solid and the melt, respectively. The difference in heat flux across the interface $z = h(t)$ results from the

latent heat released on solidification. This can be expressed by

$$\rho_s L \dot{h} = -k_m \frac{\partial T}{\partial z}(h+, t) + k_s \frac{\partial T}{\partial z}(h-, t) \quad (6)$$

where a dot denotes a time derivative and L is the latent heat of fusion. Conservation of solute requires

$$C(h+, t) \dot{h} = -D \frac{\partial C}{\partial z}(h+, t) \quad (7)$$

Further, with the assumption that the solidification kinetics can be described by equilibrium thermodynamics, we require that the temperature and composition at the interface lie on the liquidus, which we approximate by the linear relationship

$$T = T_L(C) = -\Gamma C \quad (8)$$

This tacitly assumes, merely for convenience, that the liquidus temperature is 0 at zero composition, as it is for all aqueous solutions.

The system of equations and boundary conditions represented by equations (1)–(8) admits a similarity solution, with variable

$$\eta = \frac{1}{2}(Dt)^{-1/2} z \quad (9)$$

in which the interface has position

$$h = 2\lambda(Dt)^{1/2} \equiv \gamma_C^{1/2} \quad (10, 10')$$

This similarity form of solution can be shown to be the only solution of the equations. The temperature and concentration fields, determined from equations (1)–(5), can then be expressed in terms of error functions¹⁷ as

$$T(z, t) = T_B + (T_a - T_B) \text{erf}(\varepsilon_s \eta) / \text{erf}(\varepsilon_s \lambda) \quad (z < h) \quad (11)$$

$$T(z, t) = T_\infty + (T_a - T_\infty) \frac{\text{erfc}(\varepsilon_m \eta)}{\text{erfc}(\varepsilon_m \lambda)} \quad (z > h) \quad (12)$$

$$C(z, t) = C_0 + (C_a - C_0) \frac{\text{erfc} \eta}{\text{erfc} \lambda} \quad (z > h) \quad (13)$$

where T_a and C_a are the temperature and concentration of the melt at the interface. There are two parameters $\varepsilon_{s,m} = (D/\kappa_{s,m})^{1/2}$, where $\kappa = k/\rho C_p$ is the thermal diffusivity. The interface conditions in equations (6)–(8) are now used to determine the eigenvalue λ from

$$\Gamma C_0 \left[\frac{F(\lambda)}{1 - F(\lambda)} \right] \left[\frac{\rho_m C_{p,m}}{F(\varepsilon_m \lambda)} + \frac{\rho_s C_{p,s}}{G(\varepsilon_s \lambda)} \right] = \frac{\rho_s C_{p,s} T_1}{G(\varepsilon_s \lambda)} - \frac{\rho_m C_{p,m} T_0}{F(\varepsilon_m \lambda)} - \rho_s L \quad (14)$$

where

$$F(x) = \pi^{1/2} x e^{x^2} \text{erfc } x, \quad G(x) = \pi^{1/2} x e^{x^2} \text{erf } x$$

$$T_1 = T_L(C_0) - T_B, \quad T_0 = T_\infty - T_L(C_0)$$

The solute diffusivity D is typically much smaller than the thermal diffusivity κ . In this case, $\varepsilon \ll 1$, and in the limit $\varepsilon \rightarrow 0$ equation (14) reduces to the simpler expression

$$F(\lambda) \approx [1 + (\Gamma C_0 / T_1)]^{-1} = (C_a - C_0) / C_0 \quad (15, 15')$$

Because F increases monotonically from 0 to 1, equation (15) has a unique solution of λ of order unity. This confirms the scaling used in equation (10) and indicates that the growth of the flat interface is controlled by solute diffusion. Note, however, that when C_0 is small (of the order of $\varepsilon^2 = D/\kappa$, with κ_s/κ_m of the order of unity) the growth is no longer controlled by solute diffusion but rather by thermal diffusion. This is indicated by the fact that in this event $\lambda = O(\varepsilon^{-1})$ and so $h = O(\kappa t)^{1/2}$. It would be interesting to carry out an experimental verification of the validity of equation (14).

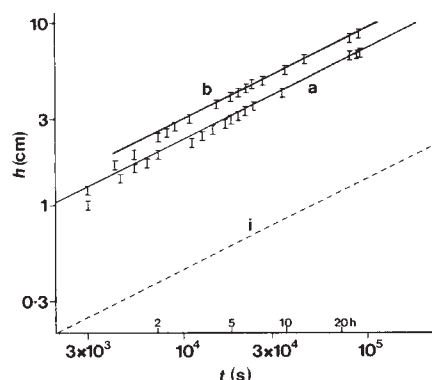


Fig. 2 Experimental and theoretical data on the height of the block as a function of time. Line a, an experiment using a Na_2CO_3 solution at eutectic composition with $T_B = -5.7^\circ\text{C}$, $T_\infty = 17.3^\circ\text{C}$. Line b is the representative experiment R with parameters in the morphologically unstable regime. Line i is the prediction of the theoretical model whereas line i is the predicted height of a flat interface at the same conditions.

Morphological instability

If the undercooling, $T_L - T_B$, is too large, solidification proceeds too quickly and the solidification interface is no longer planar. It grows in a spatially irregular manner, which will be described photographically in more detail below. The initial break-up of the planar interface is often called morphological instability. A useful physical description of the instability is discussed at length by Langer¹⁸. Detailed calculations¹⁹ indicate that, on the assumption that surface tension effects are negligible, instability occurs when

$$k_s \frac{\partial T}{\partial z}(h-, t) + k_m \frac{\partial T}{\partial z}(h+, t) = -(k_s + k_m) \Gamma \frac{\partial C}{\partial z}(h+, t) \quad (16)$$

If thermal conductivity in the solid is neglected ($k_s = 0$), equation (16) becomes equivalent to the statement that instability arises whenever the predicted values of temperature and composition on the melt side of the interface lie beneath the liquidus (where the system should be partially solid). The diffusion of heat in the solid away from the interface somewhat alters the onset of instability. Surface tension does not alter equation (16) significantly in most practical situations, so we may use it as a good approximate criterion and express it in terms of our similarity solution as

$$T_0 < \Gamma C_0 \left[\frac{F(\lambda)}{1 - F(\lambda)} \right] \left[\frac{k_m + k_s}{2k_m} \frac{F(\epsilon_m \lambda)}{\epsilon_m^2 F(\lambda)} - 1 \right] - \frac{1}{2} \frac{\rho_s}{\rho_m} \frac{L}{C_{pm}} F(\epsilon_m \lambda) \quad (17)$$

For given values of T_∞ and C_0 the value of T_B at which supercooling first occurs is found by solving equation (17) simultaneously with equation (14). The critical value of T_B is plotted as a function of C_0 for two values of T_∞ in Fig. 1a.

Experiments

We performed 12 experiments with various solutes at different values of T_B and C_0 in the tank described above. Temperatures were monitored continuously with thermistors, some of which were held fixed, and the ice block allowed to freeze around them. Occasional measurements of concentration were made by withdrawing small samples and analysing them with a hand-held refractometer. All the experiments were in the morphologically unstable regime.

The results of a representative experiment, denoted by R (NaNO_3 ; $T_B = -16.5^\circ\text{C}$, $T_\infty = 14.7^\circ\text{C}$ and $C_0 = 14 \text{ wt\%}$), are shown in Figs 2-4. Figure 3 shows three views of the solid ice block after it was removed from the tank. The humpy form of

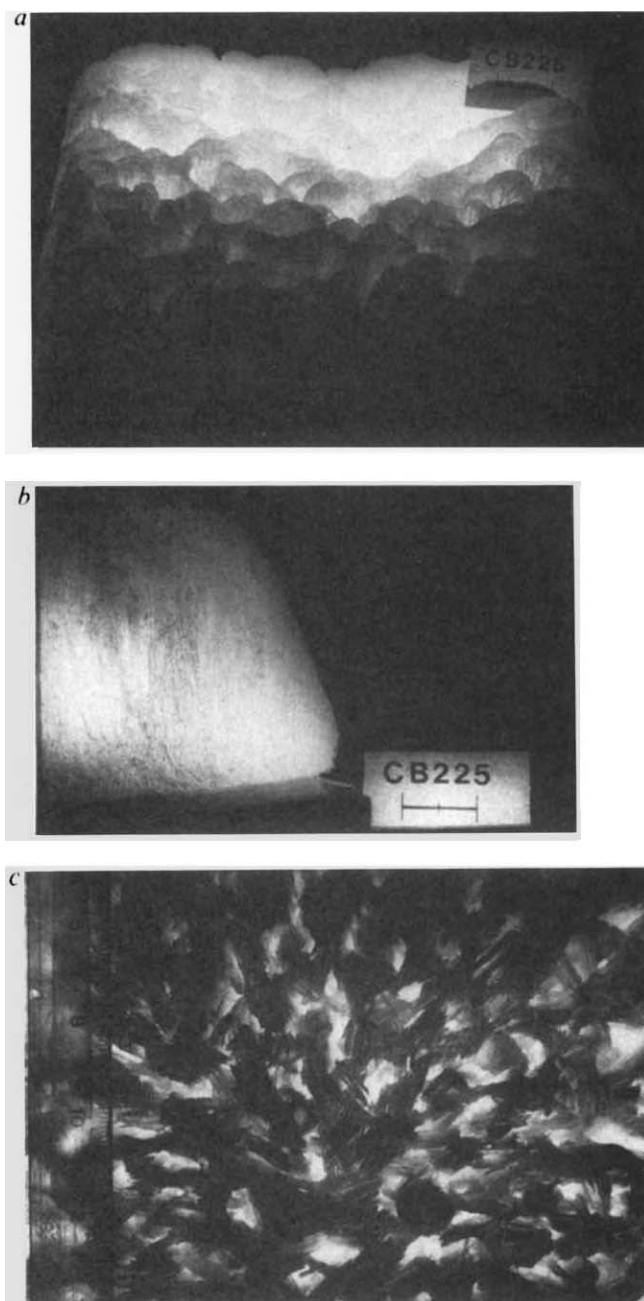


Fig. 3 a, An oblique view of the top of the ice block in experiment R showing the humpiness of the surface on a scale of centimetres and the finer scale 'facets' on a scale $< 1 \text{ mm}$. Horizontal scale, 2 cm long. b, Vertical view of the same experiment showing the lines of solid ice and the compositionally-enriched fluid. c, Close-up view looking vertically downwards onto the top of the block in the same experiment. Note the very fine-scale irregularities.

the surface (Fig. 3a) and the small-scale facets (Fig. 3c) depict clearly that the system is in the morphologically unstable regime. Figure 2 presents the (maximum) height of the block as a function of time and compares this with the result obtained by solving equations (10) and (14). The difference between the two indicates that the contortions of a morphologically unstable surface allows it to grow much more rapidly than if it were to remain stable. Figure 4 presents experimental and theoretical data on temperature and concentration at various times. There is considerable difference between the experimental results and those given by equations (11)-(14). The generality of this disagreement is confirmed in Fig. 5 which compares the experimental values of γ with those given by equations (11)-(14) for varying initial C_0 at fixed T_B and varying T_1 at fixed C_0 .

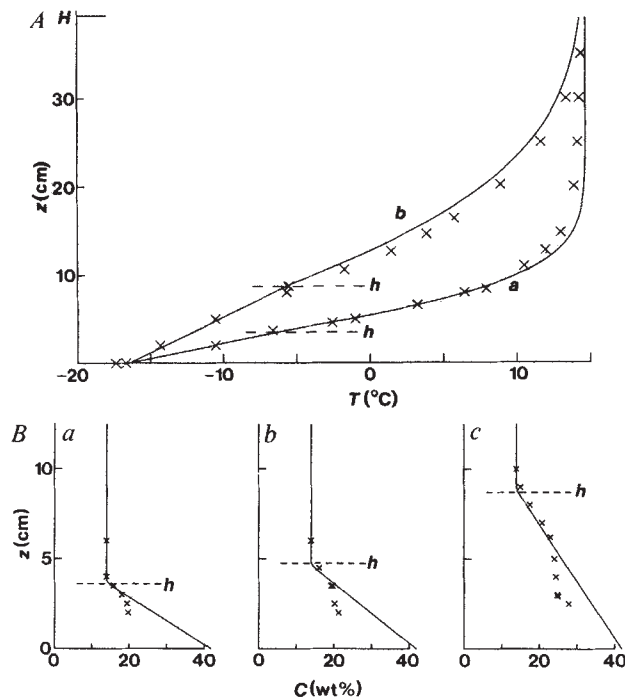


Fig. 4 A, Temperature profiles as a function of depth for experiment R at a, 3 h 40 min and b, 25 h 35 min. Crosses, data points; dashed lines, observed heights, h , of the ice block. Solid curves, temperatures predicted by the model of the mush. B, Composition profiles as a function of depth for experiment R at a, 3 h 20 min; b, 7 h 0 min; c, 25 h 20 min. Crosses, NaNO_3 concentration in fluid withdrawn by means of inserting a thin sampling tube; dashed lines, observed heights, h , of the ice block; solid curves, concentrations predicted by the model of the mush.

Theoretical model

The experimental observations described here suggest that a mixed phase of solid and melt, sometimes called a mush phase, is formed when the flat interface is morphologically unstable. Equations describing the evolution of a mush have been developed previously^{21,22} and similarity solutions have been found²¹. These equations relate bulk properties of the mush, averaged over a length scale larger than the typical crystal spacing. Thus, when a melt is cooled from below and crystallizes in such a way that relatively heavy fluid is released, the averaged equations predict convective stability of the interstitial fluid. The large pore size in the ice-blocks, however, suggested that compositional convection resulting in solute redistribution had occurred on a sub-pore scale. This would render the previous models inapplicable to the present situation. However, a simple predictive model of the growth of the mush (when $T_B > T_E$) can be developed, based on the controlling nature of thermal diffusion and the appropriate use of bulk conservation relationships.

We assume, for simplicity, that the densities of the solid and melt are the same. The mush extends from the cooled boundary to some position $z = h(t)$ and we assume the solid fraction, ϕ , within the mush to be constant. In reality, ϕ varies with height, but our model conserves mass only on a global scale and the constancy of ϕ is a necessary and consistent approximation. The solute concentration has uniform value C_0 throughout the melt in $z > h$ and obeys the liquidus relation described by equation (8) in the interstices of the mush. Diffusion governs the heat transfer in both the mush and the melt. Thus the thermal field is given by equations (11) and (12) with k_s and C_p replaced by the approximate expressions, whose validity are discussed by Batchelor²³

$$\bar{k} = \phi k_s + (1 - \phi) k_m \quad (18)$$

and

$$\bar{C}_p = \phi C_{ps} + (1 - \phi) C_{pm} \quad (19)$$

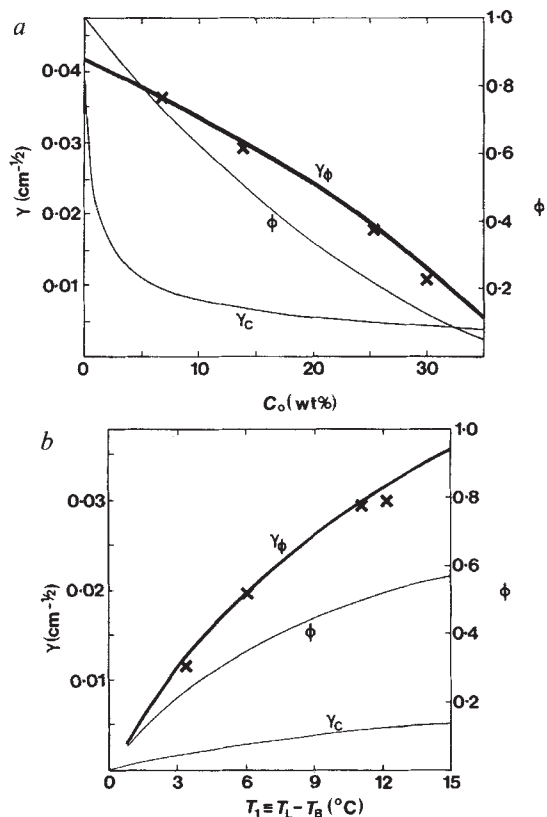


Fig. 5 a, Growth rate coefficient (γ) and solid fraction (ϕ) as functions of concentration for $T_B = -17^\circ\text{C}$ and $T_\infty = 15^\circ\text{C}$. The crosses are experimental values of λ . The heavy curve is the theoretical prediction obtained by solving equations (22) and (23). The light curve γ_c is the result of the model which assumes a flat interface and results in equation (14). b, γ and ϕ as a function of the undercooling $T_1 = T_L - T_B$ for $C_0 = 14\text{ wt\%}$ and $T_\infty = 15^\circ\text{C}$.

and T_a replaced by $T_L(C_0)$. Global conservation of solute requires that

$$(1 - \phi) \int_0^{h(t)} C dz = -\Gamma^{-1} (1 - \phi) \int_0^{h(t)} T dz = h(t) C_0 \quad (20)$$

while conservation of heat at the mush/melt interface is expressed by

$$\rho L \phi \dot{h} = -k_m \frac{\partial T}{\partial z}(h+, t) + \bar{k} \frac{\partial T}{\partial z}(h-, t) \quad (21)$$

Expressions (11) and (12) are used in equations (20) and (21) to obtain

$$\phi = \left[1 + \frac{\Gamma C_0}{T_1} H(\bar{\delta} \mu) \right]^{-1} \quad (22)$$

$$L \phi = \frac{-C_{pm} T_0}{F(\mu)} + \frac{\bar{C}_p T_1}{G(\bar{\delta} \mu)} \quad (23)$$

where

$$H(x) = [\exp(x^2) - 1]^{-1} G(x) \\ \bar{\delta} = (\kappa_m \rho \bar{C}_p / \bar{k})^{1/2}$$

and

$$\mu = \varepsilon_m \lambda$$

In terms of the new variable μ , the height of the block is given by

$$h = 2\mu (\kappa_m t)^{1/2} = \gamma_\phi t^{1/2} \quad (24, 24')$$

which confirms that the growth of the block is governed by thermal diffusion. Note, alternatively, that solute diffusion is ignored in the model and correspondingly D does not appear in equations (22), (23) or (24).

Equations (22) and (23) can be solved simultaneously for ϕ and μ . The growth rate coefficient γ_ϕ is plotted as a function of C_0 for fixed T_B in Fig. 5a and as a function of T_B for fixed C_0 in Fig. 5b. The predicted growth rate is compared with the experimental results in Figs 2 and 5 and the agreement is good.

The measured temperature profiles at two different times during experiment R compare well with the predicted temperatures (Fig. 4A) and the measured concentration profiles agree reasonably well (Fig. 4B). The disagreement between the experimental data and the theoretical model at points deep in the block may result from the difficulty of inserting a sampling tube reliably to that depth and withdrawing fluid for the concentration measurement.

The validity of the theoretical relationships was tested further by conducting some experiments with different aqueous solutions. The experimental results obtained from using NaCl and NH_4Cl , which have different liquidus relationships and diffusion

coefficients, again showed good agreement with the results of the theoretical model.

In conclusion, the experiments and related theory indicate that accurate predictions can be obtained from the use of simple conservation relationships when a melt whose composition is less than the eutectic is cooled from below. Variations of this same approach allow all six different cases, as mapped out in Table 1, to be solved. We plan to publish further details elsewhere.

We thank M. A. Hallworth for technical assistance and J. M. Wheeler, who helped with numerical calculations. We also thank D. R. Fearn, M. C. Flemings, E. J. Hinch, D. T. J. Hurle, R. C. Kerr, J. Lister, D. E. Loper and D. P. McKenzie for critical reading of the manuscript. The research was supported by grants from the BP Venture Research Unit and NERC. The paper was written while the authors were staff members at the 1984 WHOI Geophysical Fluid Dynamics Summer School, which operates under a NSF grant.

Received 3 December 1984; accepted 31 January 1985.

1. Brice, J. C. *The Growth of Crystals from Liquid* (North-Holland, Amsterdam, 1973).
2. Bardsley, W., Hurle, D. T. J. & Mullin, T. B. *Crystal Growth: a Tutorial Approach* (North-Holland, Amsterdam, 1979).
3. Laudise, R. A. *The Growth of Single Crystals* (Prentice-Hall, New York, 1972).
4. Rosenberger, F. *Fundamentals of Crystal Growth* Vol. 1 (Springer, Berlin, 1979).
5. Huppert, H. E. & Sparks, R. S. J. *A. Rev. Earth planet. Sci.* **12**, 11–37 (1984).
6. Ehlers, E. G. *The Interpretation of Geological Phase Diagrams* (Freeman, San Francisco, 1972).
7. Gubbins, D., Masters, T. G. & Jacobs, J. A. *Geophys. J. R. astr. Soc.* **59**, 57–99 (1979).
8. Loper, D. E. & Roberts, P. H. in *Stellar and Planetary Magnetism* (ed. Soward, A. M.) 297–327 (Gordon & Breach, London, 1983).
9. Flemings, M. C. *Solidification Processing* (McGraw-Hill, New York, 1974).
10. Chalmers, B. *Principles of Solidification* (Wiley, New York, 1964).
11. Hobbs, P. V. *Ice Physics* (Clarendon, Oxford, 1974).
12. Fletcher, N. H. *The Chemical Physics of Ice* (Cambridge University Press, 1970).
13. Chen, C. F. & Turner, J. S. *J. geophys. Res.* **85**, 2573–2593 (1980).
14. Huppert, H. E. & Turner, J. S. *Earth planet. Sci. Lett.* **54**, 144–152 (1981).
15. Turner, J. S., Huppert, H. E. & Sparks, R. S. J. *J. Petrol.* (in the press).
16. Carslaw, H. S. & Jaeger, J. C. *Conduction of Heat in Solids* (Oxford University Press, 1959).
17. Abramowitz, M. & Stegun, I. A. *Handbook of Mathematical Functions* (US Government Printing Office, Washington, 1964).
18. Langer, J. S. *Rev. mod. Phys.* **52**, 1–28 (1980).
19. Sekerka, R. F. *Crystal growth: an Introduction* (ed. Hartman, P.) 403–442 (North-Holland, Amsterdam, 1973).
20. Weast, R. C. (ed.) *Handbook of Chemistry and Physics* (CRC, Cleveland, 1971).
21. Worster, M. G. thesis, Univ. Cambridge (1983).
22. Hills, R. N., Loper, D. E. & Roberts, P. H. *Q. Jl Mech. appl. Math.* **36**, 505–539 (1983).
23. Batchelor, G. K. *A. Rev. Fluid Mech.* **6**, 227–255 (1974).
24. Washburn, E. W. (ed.) *International Critical Tables* (McGraw-Hill, New York, 1929).
25. Kerr, R. C. thesis, Univ. Cambridge (1984).
26. Weeks, W. F. & Ackley, S. F. *Cold Regions Res. Engng Lab. Mon.* 82–1 (1982).

Cordilleria, a newly defined Canadian microcontinent

V. E. Chamberlain & R. St J. Lambert

Department of Geology, The University of Alberta, Edmonton, Alberta, Canada T6G 2E3

Palaeomagnetic data and geological field evidence indicate that most of the western Canadian cordillera lay 1,500 km south of its present position in the early Cretaceous. We suggest that this region comprised a microcontinent, termed Cordilleria, which moved on the Kula plate and collided with the craton ~100 Myr ago to form first the Mackenzie Mountains and then the Rockies.

MODELS for the evolution of the North American cordillera have proliferated since the first attempt¹ at a plate-tectonic synthesis for western North America, based on the classical fivefold division of the cordillera in British Columbia. Earlier models concentrated on the history of palaeosubduction zones; later a long list of dextral transcurrent faults of regional significance were added^{2,3} and integration of the United States and Canadian portions of the jigsaw attempted^{3,4}. More than 50 microplates and terranes are now recognized⁵. Palaeontological^{5–10} and palaeomagnetic (refs 11–16 and E. Irving, G. J. Woodsworth and P. J. Wynne, manuscript in preparation) data clearly show the necessity for major Mesozoic transcurrent movement.

The earlier palaeomagnetic data^{11–13} indicated that elements of superterrane Terrane II (ref. 17; Fig. 1) were far travelled. More recent data (refs 14, 15 and E. Irving *et al.*, in preparation) show that Quesnellia and Stikinia, elements of Terrane I (ref. 17), are also far travelled, consistent with the palaeontological evidence^{7,8}. The most recent palaeomagnetic evidence (refs 15, 16 and E. Irving *et al.*, in preparation) suggests that both Terranes I and II have been displaced northwards by 14–20° of latitude and rotated clockwise through 45–50° with respect to the North

American craton since the early to mid-Cretaceous. These data, from two laboratories and a wide range of Triassic, Jurassic and Cretaceous rocks, show that the southern half of British Columbia was at the present-day relative latitude of California in the early Jurassic, and was probably still there in the early to mid-Cretaceous. The movement to be accommodated between Terranes I and II and the North American craton is ~1,500 km since the mid-Cretaceous. The locus of a possible innermost (eastern) transcurrent fault on which to take up this movement has remained a problem^{3,17}. Recent palaeomagnetic evidence is difficult to reconcile with the relatively small geologically observable transcurrent movements on any of the major cordilleran faults.

Transcurrent faults

The major transcurrent faults of the Canadian cordillera are shown in Fig. 2. The Denali, Queen Charlotte and Yakom-Pasayten fault systems are too far west to accommodate movement of any elements of Terrane I. The Finlay fault system (Kutcho-Kechika-Finlay-Pinchi faults) has a total offset of ~300 km¹⁸ and no true southern extension. The Fraser fault system can accommodate only ~150 km of early Tertiary move-

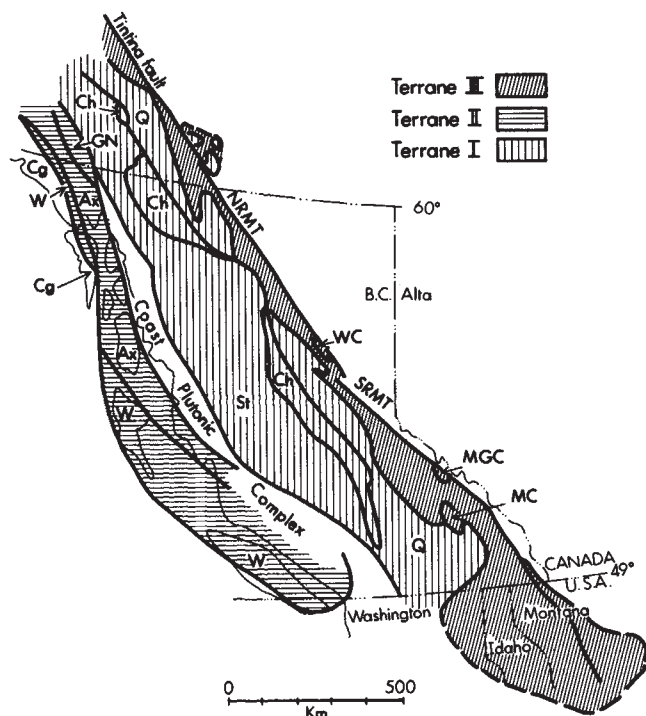


Fig. 1 Part of the Canadian cordillera, showing Terranes I and II of Monger *et al.*¹⁷ and Terrane III proposed here. Ax, Alexander terrane; Cg, Chugach terrane; Ch, Cache Creek terrane; GN, Gravina-Nutzotin terrane; Q, Quesnellia; St, Stikinia; W, Wrangellia; M, Monashee complex; MGC, Malton gneiss complex; WC, Wolverine complex; NRMT, Northern Rocky Mountain Trench; SRMT, Southern Rocky Mountain Trench.

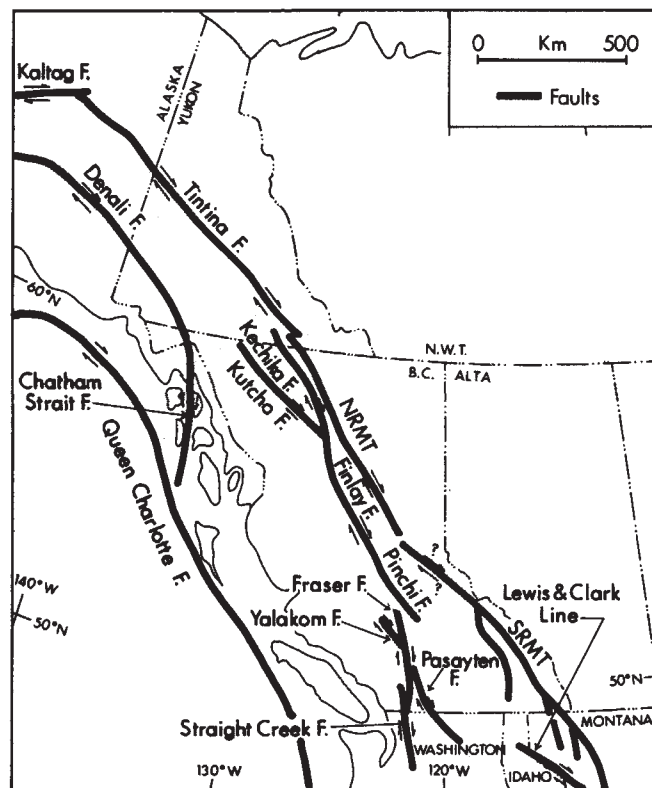


Fig. 2 Main transcurrent faults of the Canadian cordillera. None shows sufficient movement to account for all the palaeomagnetically deduced motion.

ment, which is a very small fraction of the total northward movement required of Stikinia (ref. 3 and J. W. H. Monger, personal communication). The Tintina fault and the North Rocky Mountain Trench (NRMT) fault together form a system ~1,500 km long. 450 km of late Cretaceous dextral strike slip movement has been documented on the Tintina fault¹⁹ and there may have been as much as 1,000 km of post-Palaeozoic movement on the Tintina/NRMT system (H. Gabrielse, personal communication). Thus, at least some of the northward movement of Terranes I and II may have been taken up on these faults.

Southern Rocky Mountain Trench

The Southern Rocky Mountain Trench (SRMT) is almost co-linear with the Tintina and NRMT faults, although there is a discontinuity at 54° N. Together, these three faults form one of the major linear features of the Earth's crust, with a total length of ~2,500 km. The exceptional length and linear character of the SRMT have attracted many suggestions of Mesozoic or Palaeozoic transcurrent movement along it²⁰. The large erosional feature (trench) suggests that it is the locus of a major fault, but its surface expression at the present-day stratigraphic level seems to result from thrust and dip-slip faulting²¹. First motion studies on a ~14-km deep, 4.8-mag earthquake just east of the SRMT near Valemount in British Columbia in 1978, however, showed dextral transcurrent as well as thrusting movement²². Faults frequently contain more than one type of movement²³.

Malton gneiss complex

Most stratigraphic correlations made across the SRMT are carried on thrust slices, but one major anchor seemed to be the Malton gneiss complex, which sits astride the trench at latitude 52°30'N, and has been interpreted as a single gneissic body of Proterozoic age²⁴. As such, it prohibited any large-scale post-Proterozoic movement on the SRMT.

Our investigations²⁵⁻³¹, which included major and minor element analysis of 250 gneiss samples, have shown that the Malton

gneiss complex consists of several distinct fault-bounded blocks of amphibolite facies gneiss, each of different age, geological history and petrogenesis (Fig. 3). This is supported by other authors, who describe all contacts between the gneissic basement and the Proterozoic sedimentary cover as faults, thrusts or decollements³⁰⁻³². In particular, the Malton Block to the west of the SRMT differs from the three gneissic blocks to the east of the SRMT in age, lithology, geochemistry and isotope composition²⁵⁻²⁹.

These fundamental differences between the gneisses east and west of the SRMT mean that the Malton gneiss complex is almost certainly more than one gneissic complex and can no longer be regarded as a constraint to transcurrent movement along the SRMT, nor as a tie, linking stratigraphy on one side of the SRMT to that on the other side. The gneisses do not correlate across the trench. For this reason, we have suggested that the term 'Malton gneiss complex' to describe the blocks east of SRMT should no longer be used²⁹.

Researchers who have studied the SRMT elsewhere have failed to uncover such significant differences across it, so we suggest that the Malton Block may be basement exposed in a window through the overlying late Cretaceous to Tertiary thrust sheets. Proterozoic metasediments of the Horsethief Creek Group are thrust over the Malton gneiss and across the SRMT³³. A major decollement between the Proterozoic metasediments of the Kaza Group and the underlying Malton gneiss exposed on top of the Malton Block near the head of Robina Creek (Fig. 3) may be the basal decollement of the Rockies. The gneissic blocks east of the SRMT seem to be thrust slices, but have no obvious correlatives to the west. The reason for basement exposure at this location may be a domal upwarp at the intersection of two major structural antiforms. These are the NNW-SSE trending line of domal metamorphic core complexes³⁴ and the WSW-ENE trending Jasper culmination³⁵ that brings Precambrian rocks to the surface across the entire width of the main ranges at this latitude. The metamorphic grade of the Malton gneiss complex indicates that the surface presently

exposed has moved upwards by at least 15 km since major metamorphism.

Geophysical evidence

As a result of our observations, particularly the 'window into basement' concept, and the data demonstrating the differences across the SRMT at 52°30'N, we have reviewed the geophysical evidence for the deep structure of the crust and lithosphere on either side of the SRMT. North of ~52°N, every major geophysical parameter undergoes a significant, usually abrupt, change from one side of the SRMT to the other. South of ~52°N, these geophysical changes seem to follow a north-south line up to 150 km west of the present surface expression of the SRMT and close to the line of Kootenay Lake. The structure of the lithosphere is fundamentally different east and west of this SRMT-Kootenay Lake line, shown by seismic, gravity, magnetic, electrical conductivity and heat-flow data. The seismic activity along the SRMT is moderate, with at least one earthquake of magnitude ≥ 3 every 2 yr, on an average based on the past 20 yr. In 1918, there was a 6.0-mag earthquake south of the Hugh Allan Creek block, and in 1978 a 4.8-magnitude earthquake in the Bulldog Creek area³⁶.

The thickness of both the lithosphere and the crust change markedly across this SRMT-Kootenay Lake line, both being much thinner to the west. The crust has an average thickness of 50 km east of the SRMT³⁷ and 35 km west of the SRMT³⁸⁻⁴⁰. The lithosphere west of the SRMT is also 35 km thick, its base apparently coincident with the base of the crust^{39,41}. The wider spacing of the magnetic anomalies east of this line⁴² is consistent with the lithosphere there being thicker than to the west. The abrupt change in the trend of the magnetic anomalies as this line is crossed suggests strongly a fundamental lithospheric break⁴². Details of the magnetic anomalies such as a broadening of the eastern anomalies at the SRMT suggest a cratonic or lithospheric plate margin. The magnetically quiet zones to the west of the SRMT-Kootenay Lake line and east of the Queen Charlotte fault are also typical of plate boundaries; the magnetic zone boundaries⁴² suggest that there exists a single separate plate between the Pacific plate and the Rocky Mountain Trench. This plate, we suggest, is a remnant of the Kula plate. Further evidence for the eastern boundary of this plate comes from electrical conductivity⁴³ and heat flow⁴⁴ measurements, both of which decrease sharply as the SRMT-Kootenay Lake line is crossed from west to east. Recently, a narrow region of exceptionally high electrical conductivity has been measured to the east of the SRMT under the main ranges of the Rocky Mountains near Valemount (D. I. Gough, in preparation). This is also a broad area of extremely high negative Bouguer anomaly (210 mGal) which is centred over the southern Canadian Rocky Mountains and decreases north-west and south-east around latitudes 54°N and 49°N, as well as decreasing to the north-east and south-west along the margins of the Rocky Mountain ranges⁴⁵.

The detailed structure of the lithosphere at lower crustal and mantle depths changes abruptly at the SRMT, shown by seismic and gravity measurements. A section across the SRMT at latitude 50°30'N, accommodating both kinds of data, shows the SRMT as a vertical fault at depth, but covered at upper crustal levels by continuous very low-density (2.76) material⁴⁶. It also shows a high-density (3.03) lower crustal zone to the east of the SRMT and disappearing at the trench, and a high-density (3.43) upper mantle zone, present to the east, terminating abruptly at the trench. Both are replaced to the west by 3.36 density upper mantle material. This model suggests that the SRMT is a vertical fault at depth, separating lithospheres of contrasting structures. The continuous low-density material extending laterally across the vertical fault at upper crustal levels we interpret as the Rocky Mountain thrust sheets. We suggest that the lithosphere structure here may be similar to that illustrated for the closure of the Iapetus Ocean⁴⁷. The continental crust of England is there shown being overridden by thrust sheets from Scotland, the plane of

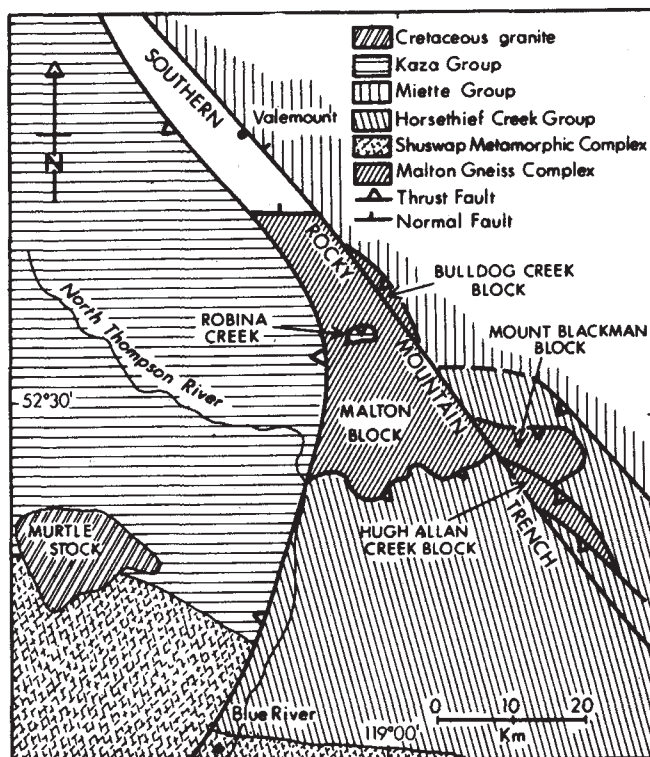


Fig. 3 The Malton gneiss complex, showing the separate blocks lying east and west of the SRMT and the overlying thrust sheets of the Proterozoic Kaza and Horsethief Creek groups.

suture dipping at a shallow angle to the north with the geophysical boundary ~100 km north of the suture in the cover rocks.

Terrane III

The Canadian cordillera has been described in terms of two superterrane, Terrane I and Terrane II, the former having amalgamated by latest Triassic-earliest Jurassic and the latter by late Jurassic¹⁷. The Omineca crystalline belt and the coast plutonic complex are described as two metamorphic and igneous welts that developed as a result of the collisions of these superterrane. The coast plutonic complex formed as a result of the accretion of Terrane II to Terrane I during the Cretaceous. The Omineca crystalline belt resulted from the mid-Jurassic collision of Terrane I with continental crustal rocks to the east¹⁷.

The rocks with which Terrane I collided during the mid-Jurassic include the north-tapering wedge of Proterozoic and Phanerozoic metasediments of the Purcell, Windermere, Hamill, Lardeau and Broadview groups, the ancient gneissic rocks of the Monashee complex (basement to the Monashee decollement⁴⁸) and possible also the Malton and Wolverine blocks. Further north, the Teslin fault is the suture zone of the mid-Jurassic collision⁴⁹, welding Terrane I to the Anvil range and Klondike schist groups.

There is no evidence of large-scale post-Jurassic north-south relative movement between Terrane I and these more easterly crustal rocks with which it collided. All the major faults in southeastern British Columbia west of the SRMT are cut by mid-Jurassic and/or Cretaceous batholiths, negating the possibility of any later major transcurrent movement. Examples occur at 125°30' W, 55°30' N; 120° W, 51°40' N; 117°30' W, 50°30' N and 117° W, 49°30' N. It seems that crustal rocks to which Terrane I became accreted in the mid-Jurassic have remained accreted to Terrane I ever since; it follows that the degree of north-south allochthoneity of these rocks is at least as great as that of Terrane I since the mid-Jurassic.

We propose the name 'Terrane III' for those crustal rocks which became accreted to Terrane I in the Jurassic; we show its approximate boundaries in Fig. 1. In general, Terrane III is bounded to the west by Terrane I and to the east by the North American craton. Figure 1 shows the Purcell thrust as the general

eastern limit of Terrane III at the present erosion level, but it is possible that it may lie further east or west. In some locations, small klippen of Terrane I occur in Terrane III and inliers of Terrane III are found in Terrane I as a result of the mid-Jurassic collision. In places, klippen of Terranes I and III have been thrust onto rocks of the North American continent but these, we suggest, were emplaced at a later time (~ 100 Myr) when Terranes I and III together collided with the North American craton.

Cordilleria

As shown earlier, the available palaeomagnetic evidence indicates that the mid-Cretaceous position of Terranes I and II was $14\text{--}20^\circ$ of latitude south of their present position relative to the North American craton and rotated $45\text{--}50^\circ$ anticlockwise of their present orientation. The Cretaceous batholiths sampled for palaeomagnetic measurements on Terrane II are from the coast plutonic complex, the welt which formed during the accretion of Terrane II to Terrane I. As Terrane III has remained accreted to Terrane I since the Jurassic, it seems probable that Terranes I, II and III were all at least loosely amalgamated and at the latitude of present-day California in early to mid-Cretaceous times.

We propose the name 'Cordilleria' for this amalgamated super-terrane or microcontinent (Fig. 4). The southern limit of Cordilleria may be at or just south of the Lewis and Clark line in the northwestern United States, although it is possible that many or all of the suspect terranes of Oregon, Washington and California were joined formerly to Cordilleria. The northern limit of Cordilleria is not clear. Wrangellia, an element of Terrane II, was amalgamated with the Peninsular terrane of southern Alaska by mid-Jurassic times, forming the Talkeetna super-terrane⁵⁰. The Peninsula and Kuskokwim terranes have been at the same latitude as Talkeetna since the mid-Jurassic⁵¹ and Talkeetna was accreted to the Nixon Fork terrane and to the Yukon-Tanana terrane (an element of Terrane I) in the mid-Cretaceous⁵⁰. This implies that a large portion of southern Alaska may have been included in the microcontinent Cordilleria.

Cretaceous subduction

Figure 5 shows one possible sequence of events for the movement of Cordilleria during the Cretaceous and here we describe a model which we believe best fits the evidence. Accurate dates are difficult to obtain because most of the available data are in the form of K-Ar ages. These tend to be minima, dating the time of final uplift and cooling of a terrane, and are also subject to problems of argon loss (and gain) and atmospheric contamination. Fitting the palaeomagnetic data into these ages causes further difficulties because palaeomagnetic blocking temperatures tend to be higher than K-Ar blocking temperatures. Further, there can be a time lag as long as 10 Myr between the onset of subduction and the K-Ar age given by a generated pluton. Present-day plate boundaries change on average every 10 Myr and every 500 km. Hence it is unreasonable to expect past plate boundaries to have been longer lasting or more continuous in type along strike.

Despite these uncertainties, it seems from the palaeomagnetic evidence that Cordilleria, relative to North America, has moved a distance of at least 1,500 km northwards since early to mid-Cretaceous. Elements of Cordilleria have rotated (possibly by internal deformation) $\leq 50^\circ$ clockwise during the same period. Relative motion between Cordilleria and the North American craton was also affected by the westward movement of the North American plate as the North Atlantic opened.

In our model, the translation and rotation of Cordilleria is effected by seafloor spreading at the Kula-Farallon and the Kula-Pacific ridges, as well as by the northward motion of the Kula-Farallon-Pacific triple junction. Not all of the movement would have occurred along the boundaries of Cordilleria; some may have taken place within the microcontinent as the loosely

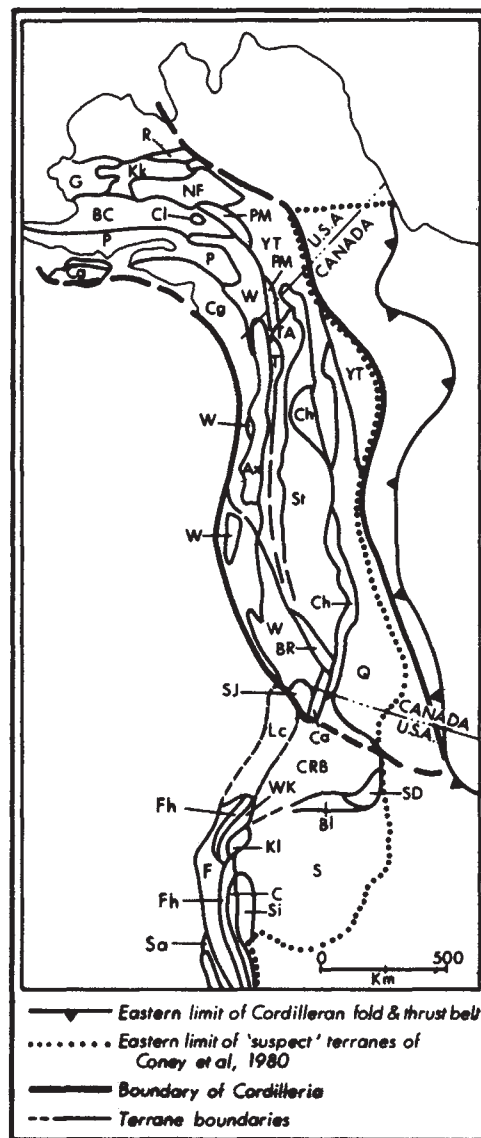
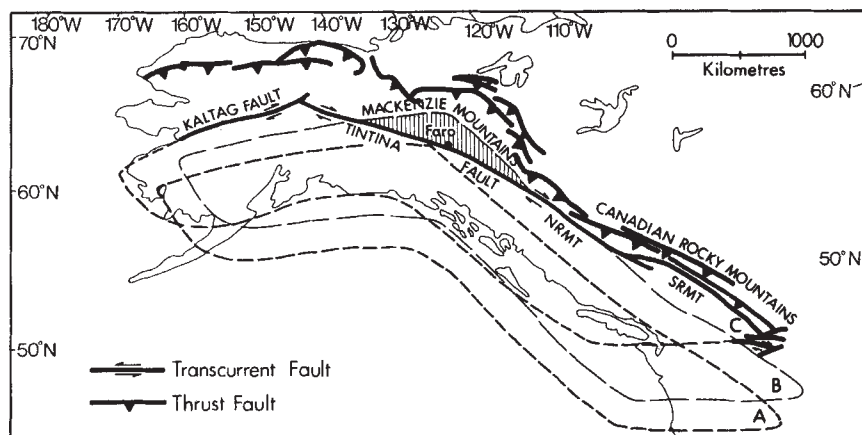


Fig. 4 The western Cordillera of North America, showing the terranes and the approximate boundaries of Cordilleria (after Coney *et al.*⁵ and Stone *et al.*⁵¹). Ax, Alexander; BC, Upper Mesozoic flysch; Bl, Blue Mountains; BR, Bridge River; C, Calaveras; Ca, northern Cascades; Cg, Chugach; Ch, Cache Creek; Cl, Chulitna; CRB, Columbia river basalts; F, Franciscan; Fh, foothills; G, Goodness; Kk, Kuskokwim; KI, Eastern Klamath Mountains; LC, Lower Cenozoic volcanic and sedimentary rocks; NF, Nixon Fork; P, Peninsular; PM, Pingston and McKinley; Q, Quesnellia; R, Ruby; S, Sonoma; SD, Seven Devils; Si, northern Sierra; SJ, San Juan; St, Stikina; T, Taku; TA, Tracy Arm; W, Wrangellia; WK, Western Klamath Mountains; YT, Yukon-Tanana; Sa, Salina.

amalgamated terranes moved relative to each other. Deformation of continental lithosphere rarely occurs on a single fault²³.

The relative motion between Cordilleria and the craton could have been taken up by transcurrent faulting, or subduction, or both. The western margin of the North American craton at this time was approximately along the line of the Tintina, NRMT and SRMT faults. Transcurrent faulting has taken place along the Tintina/NRMT fault; there is little evidence for transcurrent movement on the SRMT and no evidence of transcurrent faulting to the south. Further, the total amount of documented movement on the northern faults is insufficient to account for all of the palaeomagnetically required motion. Hence, in our model, the relative motion between Cordilleria and the North American craton is taken up largely by southward and then south-westward-dipping oblique subduction of a late Jurassic back arc

Fig. 5 Lambert equal area projection of the 'docking' of Cordilleria during the Cretaceous superimposed on the tectonic map of the present north-east Pacific (after Drummond *et al.*⁶¹). A, B, C, Possible successive positions of the Cordilleria microcontinent as it collided with the North American craton. The shape of Cordilleria is speculative, approximating its present shape throughout, although it must have changed after collision. When in position A it was probably a looser amalgamation of Terranes I, II and III. A possible protuberance which under- or overrode the North American craton at Faro is shown shaded in B. Other protuberances may also have existed, for example, one of Belt-Purcell strata (north of A) may have been thrust onto the craton when Cordilleria reached B, before transcurrent faulting moved the rest of the microcontinent into position C. Accretionary wedges forming above subduction zones along the northern and eastern margins of Cordilleria modified its shape as it moved northward. The collision at B was accompanied by tighter welding of Terranes I, II and III and a further generation of intrusives between Terranes I and II as subduction jumped westward. B represents a restoration of the 450 km of transcurrent motion on the Tintina fault and shows that the dextral transcurrent movement (B-C) would have been accompanied by an anticlockwise rotation of Cordilleria, thereby causing convergence in the Canadian Rocky Mountain region. Since then, parts of Terrane II, the Pacific rim terranes and parts of southern Alaska have moved northwestward and westward on the Denali and other faults, further modifying the shape of Cordilleria.



basin beneath the northeastern margin of Cordilleria. The Cretaceous granitoids of the Omineca belt (with their relatively high strontium isotope initial ratio) could have been generated by subduction-related heating of the downgoing slab and the overlying relatively old crustal rocks of Terrane III. Volcanoes above these plutons could account for the first appearance of bentonites in the Alberta basin at ~100 Myr⁵². Subduction, besides taking up much of the palaeomagnetically required motion, would also provide a mechanism for Cretaceous (as opposed to Jurassic) thrusting and an accommodation at depth for the ~200 km of crustal shortening apparent in the southern Canadian Rocky Mountain thrust and fold belt⁵³. An accretionary wedge building in the trench above this subduction zone produced the Mesozoic flysch terranes of southern Alaska⁵⁰, and further south it could have accumulated some of the 1×10^6 km³ of material that has been eroded from the Omineca crystalline belt since the mid-Jurassic. If half of the detritus eroded during this unroofing of the Shuswap complex was later thrust onto the North American continent (under the Rocky Mountain thrust sheets) it could account for some of the ~20 km of very low-density material there⁴⁶ and for the location of the high negative Bouguer anomaly⁴⁵. Graphite or pyritiferous black shales in this sedimentary pile could account for the very high electrical conductivity measured in this region (D. I. Gough, in preparation).

The first evidence of collision of Cordilleria with the North American craton seems to be the ~95 Myr plutons of the Mackenzie Mountains in the Yukon⁵⁴. These plutons include S-type granites⁵⁵ consistent with generation by crustal thickening as the north-east corner of Cordilleria collided with the craton (Fig. 5, positions A-B). This is the only region where Cretaceous intrusive rocks and 'suspect' terranes occur east of the Tintina-Rocky Mountain Trench fault line. Certainly there has been obduction in this area⁵⁰ and a collision near Faro would be consistent with these facts as well as with the general shape of the Mackenzie Mountain thrust belt (Figs 5 and 6a).

Collision may also have occurred at about the same time along other protuberances of the Cretaceous North American craton; for example, the thrust slices of gneissic rocks east of the SRMT near Valemount and/or the Belt and Purcell groups in the northern United States (Fig. 6b, c). Extensions of Cordilleria north and south of point B in Fig. 5 may have broken off as the microcontinent collided with the North American craton at Idaho, the northern extension now forming the Belt group of the northwestern United States and the southern extension the Seven Devils arc and Blue Mountain terranes of Oregon and possibly even some of the suspect terranes of western California (Fig. 4).

The terranes of Cordilleria telescoped during collision and plate motions changed afterwards. The subduction zone jumped westward and changed polarity, initiating the late Cretaceous-early Tertiary plutonism of the coast ranges, and effecting a final suturing of Terranes I and II. Cordilleria was constrained to move northwestward and to rotate anticlockwise along the newly formed Tintina and NRMT transcurrent faults (Fig. 5, positions B-C). Transcurrent movement may have extended to the SRMT fault as well, but examination of Fig. 5 suggests that a small wedge-shaped basinal area along the SRMT would have closed slowly as Cordilleria rotated anticlockwise. This motion would rotate the southern end of the microcontinent over the craton, pushing a miogeosynclinal wedge of thrust sheets in front of it (Fig. 6b). Such a rotation would produce maximum (~200 km⁵³) overthrusting at the southern end of the Canadian Rockies and minimum (~50 km⁵⁶) overthrusting at the northern end. Meanwhile, the final collision of the northern end of Cordilleria with northern Alaska caused the ~200 km of overriding seen there⁵¹ and the cataclasis of large portions of the northern Yukon-Tanana terrane. South of Cordilleria, the anticlockwise rotation may have been taken up by oblique northeasterly subduction of the oceanic part of the Kula plate under the North American craton producing the Idaho batholith (Fig. 6c). The radiometric ages obtained on the Idaho batholith⁵⁷ and the Valhalla dome⁵⁸ as well as the youngest ages obtained on the Proterozoic and Archaean metamorphic units along the SRMT²⁸ are mainly in the range 60-90 Myr. We suggest that these are intrusive or cooling ages marking the time of final suture of Cordilleria with North America, at ~80 Myr, a time when the rate of opening of the North Atlantic reached a maximum⁵⁹.

During the northward movement of Cordilleria and during and after its time of collision with the North American craton, further readjustments of the individual terranes took place; the Fraser, Denali and Queen Charlotte faults all record post-Cretaceous transcurrent movement. Evidence of late Tertiary extension is seen in normal faulting on many of the southern Cordilleria faults, including the SRMT.

The rate of northward movement of Cordilleria, relative to the North American craton, is difficult to estimate because of the problem mentioned above of obtaining exact ages on intrusive rocks. The most likely dates are 120-110 Myr for a position ~1,500 km south of present and 105-100 Myr for position A (Fig. 5) which would require ~1,000 km of northward movement in 10-15 Myr at an average rate of 100-150 mm yr⁻¹. (The remaining 500 km of northward movement were taken up by transcurrent movement in the north along the Tintina fault, and by oblique subduction and thrusting in the south.)

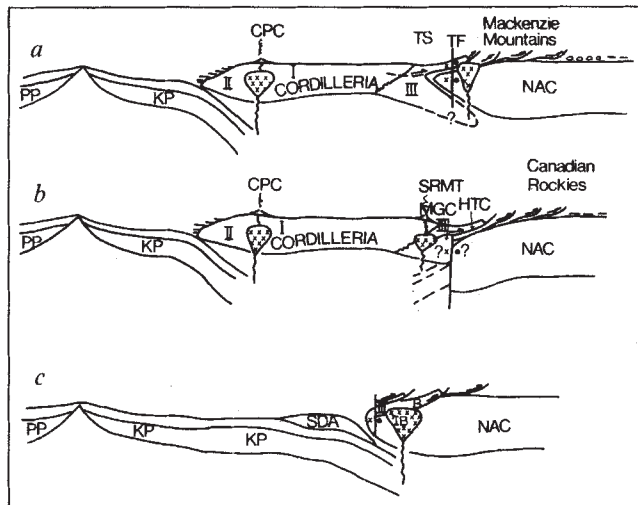


Fig. 6 Cross-sections at ~80 Myr drawn at three different latitudes. *a*, Section at the approximate latitude of Faro, Yukon, 62° N. *b*, Section at the latitude of the Malton gneiss complex near Valemount, BC 52°30' N. *c*, Section at the approximate latitude of Salmon, Idaho 45° N. B, Belt group; CPC, coast plutonic complex; HtC, Horsethief Creek group; IB, Idaho batholith; KP, Kula plate; MGC, Malton gneiss complex; NAC, North American craton; PP, Pacific plate; SDA, Seven Devils arc; SRMT, Southern Rocky Mountain Trench; TF, Tintina fault; TS, Teslin suture. I, II and III indicate Terranes I, II and III. *a*, Westward-dipping subduction under Cordilleria during the early Cretaceous has closed a late Jurassic back arc basin between the North American craton and Cordilleria, causing a collision at ~100 Myr. The upper crust of Cordilleria was thrust onto the craton, carrying with it Terrane I ophiolites (Nisutlin, Simpson and Anvil allochthons) which had previously been obducted onto Terrane III during the mid-Jurassic collision of Terranes I and III along the Teslin suture. Crustal thickening due to the Cretaceous collision generated the S-type granites which invade the thrust sheets of the Mackenzie Mountains. After collision, motions changed, initiating dextral transcurrent movement on the Tintina fault, and subduction jumped westward, generating the granitoids of the coast plutonic complex and causing a final suturing of Terrane II to Terrane I. *b*, Early Cretaceous westward-dipping subduction which generated granitoids along the eastern edge of Cordilleria (Murtle stock at this latitude) ceased with the collision of Cordilleria and the North American craton along the line of the SRMT. The main collision at this latitude was later than in *a* because it was caused by the anticlockwise rotation of Cordilleria as it moved along the Tintina fault. During the collision, supracrustals from Cordilleria (Horsethief Creek, Kaza) were thrust onto the craton covering the suture (SRMT) between the basement blocks. Subduction jumped westward as in *a*. *c*, The first collision occurred at about the same time as in *a*, thrusting supracrustals from Cordilleria (belt group) onto the North American craton. After the collision, Cordilleria, carried on the Kula plate, moved northwestward ~450 km, rotating anticlockwise. The northwestward translation moved the southern margin of Cordilleria into a position north of Salmon; the Seven Devils arc may be a remnant of Cordilleria, broken off as the main body of the microcontinent moved northwestward. The anticlockwise rotation caused the southern end of the Kula plate to swing northeastward, the motion being taken up by its subduction under the North American plate, generating the Idaho batholith.

We propose that Cordilleria was carried northwards on the Kula plate as it moved away from the Kula–Farallon rift opening to the south. The calculated rate of movement of the Kula with respect to the American plate, obtained by extrapolating magnetic measurements back to the beginning of the Cenozoic, is 120 mm yr^{-1} northwards⁶⁰, a value quite close to the rate of movement suggested here. At the same time, the Kula–Pacific ridge was spreading and the American plate was moving westwards as the North Atlantic opened.

Conclusions

The increasing and consistent palaeomagnetic data on Mesozoic igneous rocks of south-central British Columbia places them at

the latitude of California and rotated 45° anticlockwise from their present orientation. Existing theories do not include any line of suture between the Intermontane belt and the Rockies along which Mesozoic or early Tertiary northward movement and closure could have occurred. Our investigations show that the Malton gneiss complex (52°20' N on the SRMT) is a composite feature, a window into a basement which differs on either side of the SRMT. Noting that all geophysical parameters change across the SRMT, we consider that it marks a line of weakness above a Cretaceous suture, now largely covered by the thrust sheets of the Rockies.

All the evidence can be reconciled if a newly defined microcontinent, Cordilleria, corresponding to most of British Columbia and part of the Yukon and Alaska and perhaps parts of Oregon and California, existed. This microcontinent must have lain on the Kula plate and west of a subduction zone, along which convergence occurred during the late Jurassic and early Cretaceous, closing a back-arc basin and bringing Cordilleria northward as well as towards the craton. The latter overrode the trench and was struck by Cordilleria in the South Yukon, causing the Mackenzie Mountains at ~100 Myr. Closure of the remaining oceanic gap between the craton and Cordilleria caused Rocky Mountain thrusting over the next 40 Myr. In our model, the Tintina–Rocky Mountain Trench system marks the line of the former trench system and its associated dextral faults. As the convergence was oblique, however, and much of the 1,500–2,000 km displacement was taken up by subduction, equivalent amounts of transcurrent movement will not have occurred along all parts of the suture. Any transcurrent faults generated would tend to grow at one or both ends (as with the present day San Andreas) with maximum relative faulting at the point of first collision.

We thank E. Irving, D. I. Gough, W. S. McKerrow and many other colleagues for helpful discussions. Financial support was provided by the Canadian National Science and Engineering Research Council and the Federal Department of Energy Mines and Resources.

Received 10 October 1984; accepted 4 February 1985.

- Monger, J. W. H., Souther, J. G. & Gabrielse, H. *Am. J. Sci.* **272**, 577–602 (1972).
- Monger, J. W. H. & Price, R. A. *Can. J. Earth Sci.* **16**, 770–791 (1979).
- Davis, G. A., Monger, J. W. H. & Burchfiel, B. C. *Pacif. coast Paleogeogr. Symp.* **2**, 1–32 (1978).
- Dickinson, W. R. *Can. J. Earth Sci.* **13**, 1268–1287 (1976).
- Coney, P. J., Jones, D. L. & Monger, J. W. H. *Nature* **288**, 329–333 (1980).
- Tozer, E. T. *Geol. Surv. Can. econ. Rep.* **1**, 633–640 (1970).
- Monger, J. W. H. & Ross, C. A. *Can. J. Earth Sci.* **8**, 259–278 (1971).
- Tipper, H. W. *Can. J. Earth Sci.* **18**, 1788–1792 (1981).
- Siberling, N. J. & Jones, D. L. *Geol. Ass. Can. Abstr.* **8**, A62 (1983).
- Tozer, E. T. *Geol. Rdsch.* **71**, 1077–1104 (1982).
- Packer, D. R. & Stone, D. B. *Can. J. Earth Sci.* **11**, 976–997 (1974).
- Hillhouse, J. W. *Can. J. Earth Sci.* **14**, 2578–2592 (1977).
- Van der Voo, R., Jones, M., Gromme, C. S., Eberlein, G. D. & Churkin, M. *J. geophys. Res.* **85**, 5281–5296 (1980).
- Irving, E., Monger, J. W. H. & Yole, R. W. *Geol. Ass. Can. spec. Pap.* **20**, 441–455 (1981).
- Irving, E., Wynne, P. J. & Woodsworth, G. *J. Geol. Ass. Can. Abstr.* **8**, A35 (1983).
- Symons, D. T. A. *Geol. Ass. Can. Abstr.* **9**, 109 (1984).
- Monger, J. W. H., Price, R. A. & Tempelman-Kluit, D. *J. Geology* **10**, 70–75 (1982).
- Gabrielse, H. & Dodds, C. J. *Geol. Ass. Can. Abstr.* **2**, 19 (1977).
- Tempelman-Kluit, D. *J. Geol. Surv. Can. Pap.* **77-1A**, 223–227 (1977).
- Leech, G. B. *Geol. Surv. Can. Pap.* **66-14**, 307–329 (1966).
- Murphy, D. C. *Geol. Surv. Can. Pap.* **84-1A**, 627–630 (1984).
- Rogers, G. C., Ellis, R. M. & Hasegawa, H. S. *Bull. seism. Soc. Am.* **70**, 1771–1786 (1980).
- McKenzie, D. & Jackson, J. *Earth planet. Sci. Lett.* **65**, 182–202 (1983).
- Campbell, R. B. *Geol. Surv. Can. Map* **15-1967** (1968).
- Chamberlain, V. E., Lambert, R. St J., Baadsgaard, H. & Gale, N. H. *Geol. Surv. Can. Pap.* **79-1B**, 45–50 (1979).
- Chamberlain, V. E., Lambert, R. St J. & Holland, J. G. *Precamb. Res.* **11**, 1–9 (1980).
- Chamberlain, V. E., Lambert, R. St J. & Holland, J. G. *Precamb. Res.* **11**, 297–306 (1980).
- Chamberlain, V. E. thesis, Univ. Alberta (1983).
- Chamberlain, V. E., Lambert, R. St J. & Holland, J. G. *Can. J. Earth Sci.* (in the press).
- Morrison, M. L. *Geol. Surv. Can. Pap.* **79-1B**, 407–410 (1979).
- Oke, C. & Simony, P. S. *Geol. Surv. Can. Pap.* **81-1A**, 181–184 (1981).
- McDonough, M. R. & Simony, P. S. *Geol. Surv. Can. Pap.* **84-1A**, 99–102 (1984).
- Leonard, R. *Geol. Surv. Can. Pap.* **84-1A**, 121–127 (1984).
- Coney, P. J. *Geol. Soc. Am. Mem.* **153**, 7–31 (1980).
- Price, R. A. & Mountjoy, E. W. *Geol. Ass. Can. Spec. Pap.* **6**, 7–25 (1970).
- Ellis, R. M. & Chandra, B. *Can. J. Earth Sci.* **18**, 1708–1716 (1981).
- Mereu, R. F., Majumdar, S. C. & White, R. E. *Can. J. Earth Sci.* **14**, 196–208 (1977).
- Berry, M. J. & Forsyth, D. A. *Can. J. Earth Sci.* **8**, 788–801 (1975).
- Wickens, A. J. *Can. J. Earth Sci.* **14**, 1100–1115 (1977).
- Spence, G. D., Clowes, R. M. & Ellis, R. M. *Bull. seism. Soc. Am.* **67**, 543–546 (1977).
- Wickens, A. J. & Pec, K. *Bull. seism. Soc. Am.* **58**, 1821–1831 (1968).
- Haines, G. V., Hannaford, W. & Riddihough, R. P. *Can. J. Earth Sci.* **8**, 387–391 (1971).
- Dragnet, H. & Clarke, G. K. C. *J. Geophys.* **42**, 373–390 (1977).
- Davis, E. E. & Lewis, T. J. *Can. J. Earth Sci.* **21**, 715–726 (1984).

45. *Gravity Map Ser. No. 74-1* (Department of Energy, Mines and Resources, Canada, 1974).
46. Chandra, N. N. & Cumming, G. L. *Can. J. Earth Sci.* **9**, 1099-1109 (1972).
47. Leggett, J. K., McKerrow, W. S. & Soper, N. J. *Tectonics* **2**, 187-210 (1983).
48. Brown, R. L. & Murphy, D. C. *Can. J. Earth Sci.* **19**, 456-465 (1982).
49. Templeman-Kluit, D. J. *Geol. Surv. Can. Pap.* **79-14**, 27 (1979).
50. Csejty, B., Cos, D. P., Everts, R. C., Stricker, G. D. & Foster, H. L. *J. geophys. Res.* **87**, 3741-3754 (1982).
51. Stone, D. B., Panuska, B. C. & Packer, D. R. *J. geophys. Res.* **87**, 3697-3707 (1982).
52. Spears, D. A. & Duff, P. *Can. J. Earth Sci.* **21**, 465-476 (1984).
53. Price, R. A. in *Thrust and Nappe Tectonics* Vol. 9 (eds MacClay, K. & Price, N. J.) 427-448 (Geol. Soc. Lond. spec. Publ., 1981).
54. Anderson, R. G., Armstrong, R. L., Parrish, R. & Bowman, J. R. *Geol. Ass. Can. Abstr.* **8**, A2 (1983).
55. Anderson, R. G. *Geol. Ass. Can. Abstr.* **8**, A2 (1983).
56. Thompson, R. I. in *Thrust and Nappe Tectonics* Vol. 9 (eds MacClay, K. & Price, N. J.) 449-462 (Geol. Soc. Lond. spec. Publ., 1981).
57. Armstrong, R. L., Taubeneck, W. H. & Hales, P. O. *Bull. geol. Soc. Am.* **88**, 397-411 (1977).
58. Parrish, R. & Ryan, B. *Geol. Ass. Can. Abstr.* **8**, A53 (1983).
59. Briden, J. C., Hurley, A. M. & Smith, A. G. *J. geophys. Res.* **86**, 11631-11656 (1981).
60. Atwater, T. *Bull. geol. Soc. Am.* **81**, 3513-3536 (1970).
61. Drummond, K. J., Nishiwaki, C., Corvalan, J., Douth, H. F. & Craddock, C. *Am. Ass. Petrol. Geol. Tectonic Map* (1982).

Structural analysis of murine genes containing homoeo box sequences and their expression in embryonal carcinoma cells

Anamaris M. Colberg-Poley, Stephan D. Voss, Kamal Chowdhury & Peter Gruss

Zentrum für Molekulare Biologie der Universität Heidelberg, Im Neuenheimer Feld 364, D-6900 Heidelberg, FRG

The presence of homoeo box sequences in the genomes of vertebrates has suggested that these metazoans possess the equivalent of the homoeotic genes that have a key role in regulating the development of the fruitfly Drosophila melanogaster. We report here that a novel murine homoeo box-containing gene is expressed in embryonal carcinoma stem cells. Transcripts of the sequences that flank the homoeo box of this gene are found in these cells before and after induced differentiation, whereas a specific transcript that seems to contain the homoeo sequence is only present after differentiation.

THE establishment of embryonal carcinoma (EC) cell lines has facilitated greatly the study of early stages in mouse embryogenesis. EC cells are derived from the pluripotent stem cells of teratocarcinomas¹ and possess many properties of ectodermal cells of the inner cell mass². Advantages of these cells include the existence of various EC lines capable of differentiating into several tissue types *in vitro*, including extraembryonal endoderm, as well as their parallel use for *in vivo* studies². One of these EC stem-cell lines, F9, was isolated from a transplanted male embryo³ and expresses many antigens that are ascribed generally to embryonal carcinoma cells and mouse developing embryos⁴⁻⁷. F9 cells, unlike many other EC stem-cell lines, give rise spontaneously to differentiated cells at a very low rate⁸, although differentiation can be induced by retinoic acid in the presence or absence of dibutyl cyclic (c)AMP⁹⁻¹¹ or by transfection with *c-fos* proto-oncogenes¹². The induced phenotypic alterations of retinoic acid-differentiated F9 cells are reflected in the production of markers characteristic of either visceral endoderm¹⁰⁻¹³ or parietal endoderm^{9,13-15}, depending on the culture conditions. These extraembryonal tissues derive before implantation, during embryonal development, from a common stem-cell population, primitive endoderm.

Until recently the isolation of mouse genes expressed during development relied mainly on the identification of genetic information which is expressed discriminately in undifferentiated or differentiated EC cells^{16,17}. In turn, we have used an alternative approach to attempt to isolate genes expressed during mouse development. This approach is based on the assumption that certain mouse developmental processes might also require the expression of homoeotic genes, similar to those of the fruitfly *Drosophila melanogaster*. To this end we screened mouse genomic information with *Drosophila* homoeo box sequences. These sequences (~200 base pairs, bp) are conserved in many of the developmental genes located in the antennapedia (ANT-C) and bithorax complexes of *Drosophila*¹⁸⁻²¹. These regions share a high degree of conservation not only in *Drosophila* but also with homoeo boxes found in human²², murine²³ and *Xenopus*^{24,25} DNA. Most impressive, however, is the fact that at least two of these genes have been isolated exclusively by their homology to *Drosophila* homoeo boxes and are transcribed in oocytes or during early embryogenesis in *Xenopus*^{24,25}. Using two *Drosophila* probes we have isolated murine DNA sharing

homology with previously described homoeo box sequences. This murine DNA detects the expression of sequences next to the homoeo box in EC stem cells as well as an additional RNA species that contains homoeo box sequences and accumulates during the differentiation of these cells. All these transcripts are encoded by the same DNA strand and are not detected in RNA isolated from macrophages.

Isolation of murine homoeo box sequences

At least eight *EcoRI* fragments of murine DNA hybridize with *Drosophila* homoeo box probes¹⁹. To isolate mouse homoeo box sequences, ~3 × 10⁵ members (~1 genome equivalent) of a mouse genomic library (gift of H. Lehrach) in λEMBL3 (ref. 26) were screened under conditions of reduced stringency using purified inserts of p903G, the 3' region of a cDNA clone of *Antennapedia* (*Antp*)^{18,27}, and pFS2, the 3' region of a genomic clone of *fushi tarazu* (*ftz*)²⁸. Homologies between these two probes are confined to the sequences contained in their homoeo boxes. Of the recombinant phage screened, 10 recombinants were isolated, based on their hybridization to both probes. The DNAs of representative phage, cleaved with *EcoRI*, are shown in Fig. 1a. Hybridization of *ftz* and *Antp* probes to Southern blots²⁹ of these cleaved DNAs shows isolate-specific as well as probe-specific patterns. For example, λm5 and λm12 share two common fragments that hybridize similarly to both *Antp* and *ftz* probes, but λm6 shows hybridization of three fragments with the *ftz* probe while only two fragments hybridize to the *Antp* probe. As a control we used λm9, which is isolated by hybridization to the *ftz* probe alone. Hybridization of its *EcoRI* fragments confirms that this phage DNA hybridizes exclusively to *ftz* probe.

To determine which of the various fragments in λm5 and λm6 were hybridizing exclusively to homoeo boxes and which to flanking sequences contained within the probes, smaller fragments from *ftz* and *Antp* containing almost exclusively homoeo box sequences were hybridized with the cleaved recombinant DNA (Fig. 1b). λm5 and λm6 DNA showed hybridization of one and two *EcoRI* fragments, respectively, with both probes. The additional very faint band in λm5 which hybridizes with *Antp* is a partial digestion product and has not been otherwise observed. One of these recombinants, λm6, was mapped in

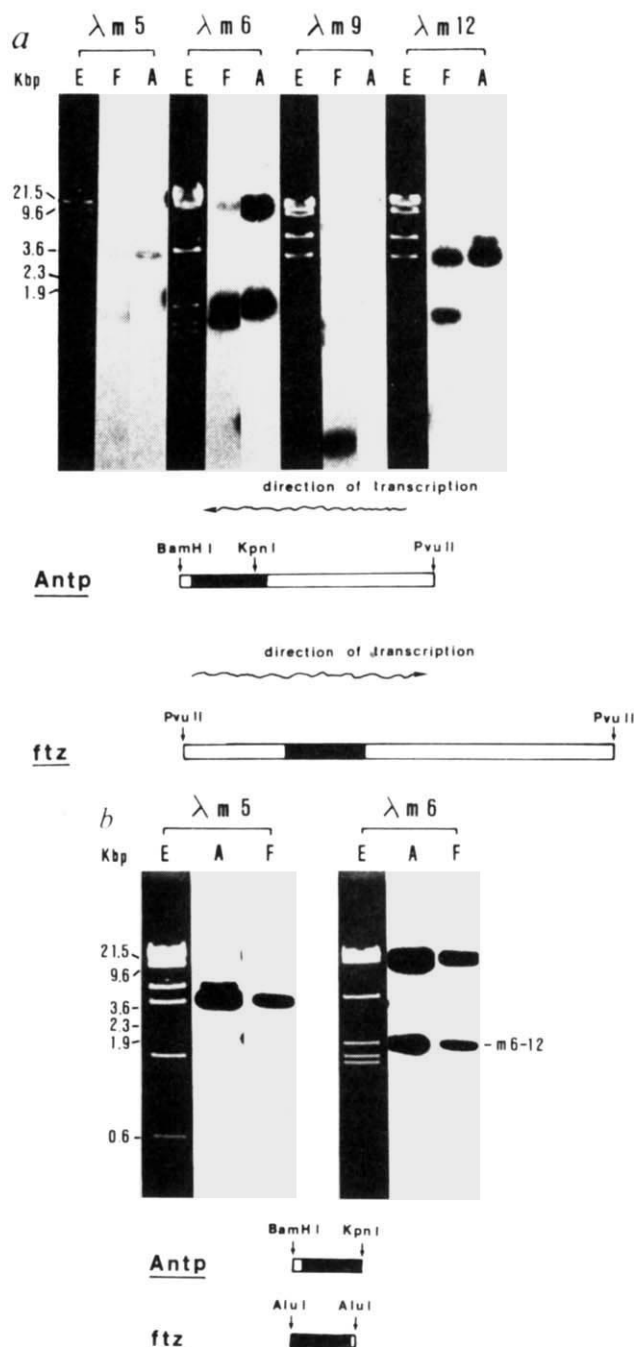


Fig. 1 Hybridization of inserts containing *Drosophila* homoeo boxes to λ mouse recombinant DNAs. *a*, Hybridization of λ mouse recombinants (λm5, 6, 9 and 12) DNAs to purified inserts of p903G (BamHI/PvuII) or pFS2 (PvuII/PvuII). Lanes E, ethidium bromide staining; F, *ftz* probe; A, *Antp* probe. *b*, Hybridization of *Drosophila* homoeo boxes to murine DNA. Phage DNAs were cleaved with *Eco*RI, separated on agarose gels (E), blotted onto nitrocellulose and hybridized with probes from *ftz* (F) or *Antp* (A). The structure of the probes and the homoeo box sequences therein (closed box) are represented beneath each panel.

Methods. Phage DNAs were isolated as described previously⁴⁰, cleaved with *Eco*RI and separated by electrophoresis in an agarose gel. Southern blots of the phage DNAs were hybridized initially with either *a*, the ³²P-labelled BamHI/PvuII fragment of p903G that contains a *Drosophila* homoeo box, then exposed, boiled in H₂O for 1 min to remove the probe and rehybridized to the homoeo box insert (PvuII/PvuII) of pFS2; or *b*, the nick-translated BamHI/KpnI insert of p903G (A) and an AluI restriction fragment of pFS2 (F). The hybridizations were performed under conditions of reduced stringency (43% formamide, 5×SSC, 5×Denhardt's, 100–250 μg ml⁻¹ denatured salmon sperm DNA, 50 mM Na₂PO₄, pH 7.0, 0.1% SDS, 37 °C) as described by McGinnis *et al.*¹⁹ with washes at 50–55 °C in 2×SSC, 0.1% SDS.

detail and, based on overlapping restriction, hybridization patterns (Fig. 2) and subsequent sequence data, found to contain two homoeo box regions lying within 10 kilobases (kb) of one another on a contiguous stretch of murine DNA.

Murine homoeo box sequences

The hybridizing 1.9-kb *Eco*RI fragment of λm6, m6-12, was isolated subsequently and subcloned into pSP65. Finer mapping of this subcloned fragment showed that the homoeo box homologous region is confined to a 206-bp *Pvu*II/*Eco*RI fragment (Fig. 2). The sequence of this region, as well as the adjacent 300-bp *Eco*RI fragment, was obtained along both strands using chain termination³⁰ and chemical modification³¹ techniques (Fig. 3a). Comparison of this sequence with the sequences of *Drosophila* (*Antp*, *ftz* and *Ubx*)¹⁹, *Xenopus* (MM3)²⁵, murine (Mo-10)²³ and human (*Hu-1* and *Hu-2*)²² homoeo boxes indicated striking, albeit varying, degrees of homology. The sequence m6 contains 83% homology to the homoeo boxes of *Hu-1* and *Hu-2*, whereas comparison with Mo-10 surprisingly showed only 78% homology; m6 is 84 and 82% homologous to *Antp* and MM3, respectively, although only 79 and 76% homologous to the other *Drosophila* homoeo boxes, *Ubx* and *ftz*, respectively.

Conceptual translation of the sequence obtained and comparison with the published homoeo box sequences of MM3, *Antp*, Mo-10, *Hu-1*, *Hu-2*, *ftz* and *Ubx* (refs 19, 22, 23, 25) are shown in Fig. 3b. The homologies in the amino-acid sequences were higher than those obtained with the nucleotide sequence because most nucleotide differences occur at the wobble position, indicating that the basis for the conservation of this region probably occurs at the protein level. The most homologous sequence showed the following homologies: to *Antp* (95%), *Xenopus* (93%), *Hu-2* (88%), *Hu-1* (85%), *Ubx* (85%) and *ftz* (80%). Surprisingly, the least homology in amino-acid sequence was seen with the mouse (Mo-10) homoeo box (72%). These sequences in the mouse seem to be the most divergent. We also observed at the end of the conserved homoeo box domain a succession of glutamic acid residues followed by a stop codon. As this nucleotide sequence was obtained along both strands using both sequencing methods, we are confident of its validity. A similar succession of acidic residues followed by an in-frame stop codon was observed by Müller *et al.*²⁵ in *Xenopus* and is also present in the sequence of *Hu-2* (ref. 22).

Expression of murine DNA

To determine whether the probe used for expression studies contained repetitive DNA, m6-12 DNA was hybridized with genomic blots of murine DNA (Fig. 4). DNA from F9 and another EC cell line, P19 (ref. 32), was cleaved with *Eco*RI, *Hind*III or *Kpn*I. Hybridization of m6-12 DNA with the cleaved DNAs under stringent conditions revealed hybridization to a single band in each case, showing that the sequences in m6-12 DNA detect a single-copy region. The expression of the murine DNA sequences containing homoeo box sequences and those adjacent to them was examined in three teratocarcinoma stem-cell lines. RNA was isolated from EK³³ cells and from various EC cell lines (PC13, P19 and F9 cells) and hybridized to plasmid DNA containing m6-12 (Fig. 5a). Two predominant bands appear in the RNAs from EK, PC13, P19 and F9 stem cells although in greatly varying amounts. The size of the larger and less abundant RNA is ~2.9 kb and the smaller RNA 1.5 kb. The quantities of poly(A⁺) RNA from PC13, P19 and F9 blotted were corroborated by hybridization to β-actin. We observed quantitative differences in the amount of the 2.9- and 1.5-kb RNAs in the various stem-cell lines. There is 3–6 times more of the 1.5-kb messenger RNA in the F9 cells than in PC13 and P19 cells and ~12 times more of this transcript in F9 cells than in EK cells. The larger 2.9-kb mRNA is also more abundant in F9 cells compared with P19, PC13 and EK cells (an ~3.5 to 5-fold difference). These RNAs are ~10 times less abundant in F9 cells than β-actin RNA. These differences were normalized

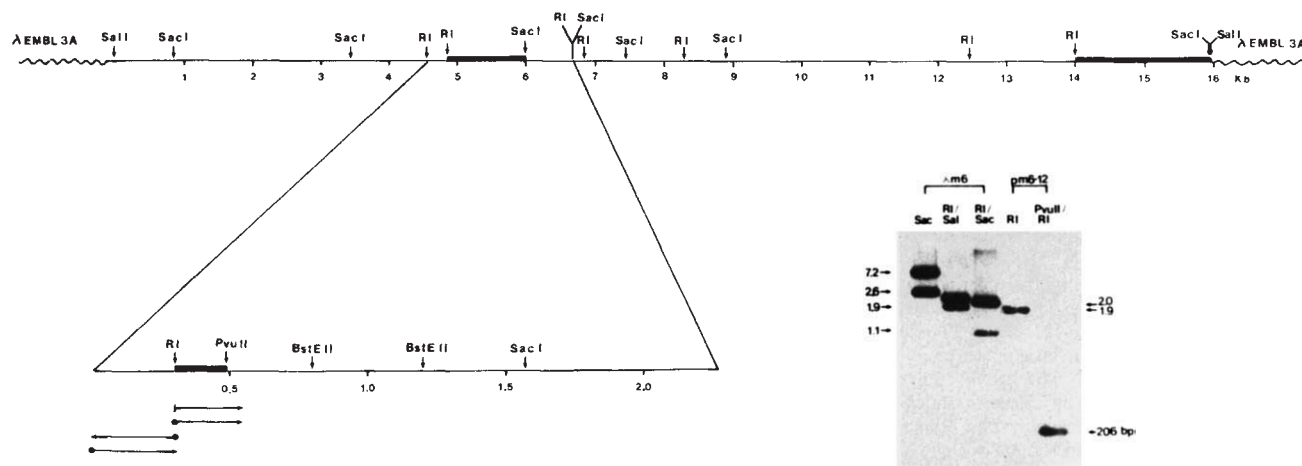


Fig. 2 Restriction mapping and subcloning of fragments contained in λ clone m6. The recombinant phage DNA contains *MboI* partially digested (~16 kb) murine liver DNA cloned into *BamHI* sites of λ EMBL3A (ref. 26). The figure localizes the homoeo box-homologous regions to 1.3-kb *EcoRI/SacI* and 2.0-kb *EcoRI/Sall* fragments (closed boxes).

Methods. Hybridizations used the 600-bp *Antp* cDNA fragment (see Fig. 1a) as probe. A 1.9-kb *EcoRI* fragment containing the 1.3-kb *SacI/EcoRI* homoeo box-hybridizing fragment was subcloned into pSP65 to give pm6-12. The homoeo box homology was further confined in pm6-12 to a 206-bp *PvuII/EcoRI* subfragment (see insert). The *BstEII/EcoRI* fragment containing the *PvuII* site was purified subsequently, 3' end labelled at the *EcoRI* site with Klenow fragment and sequenced according to the method of Maxam and Gilbert³¹. This sequence was confirmed along the opposite strand by the method of Sanger *et al.*³⁰ using a subclone of the 1.3-kb *SacI/EcoRI* fragment, mp11/m6-12. The adjacent 300-bp *EcoRI* fragment was cloned in both orientations in M13mp11 to obtain sequence information along both strands.

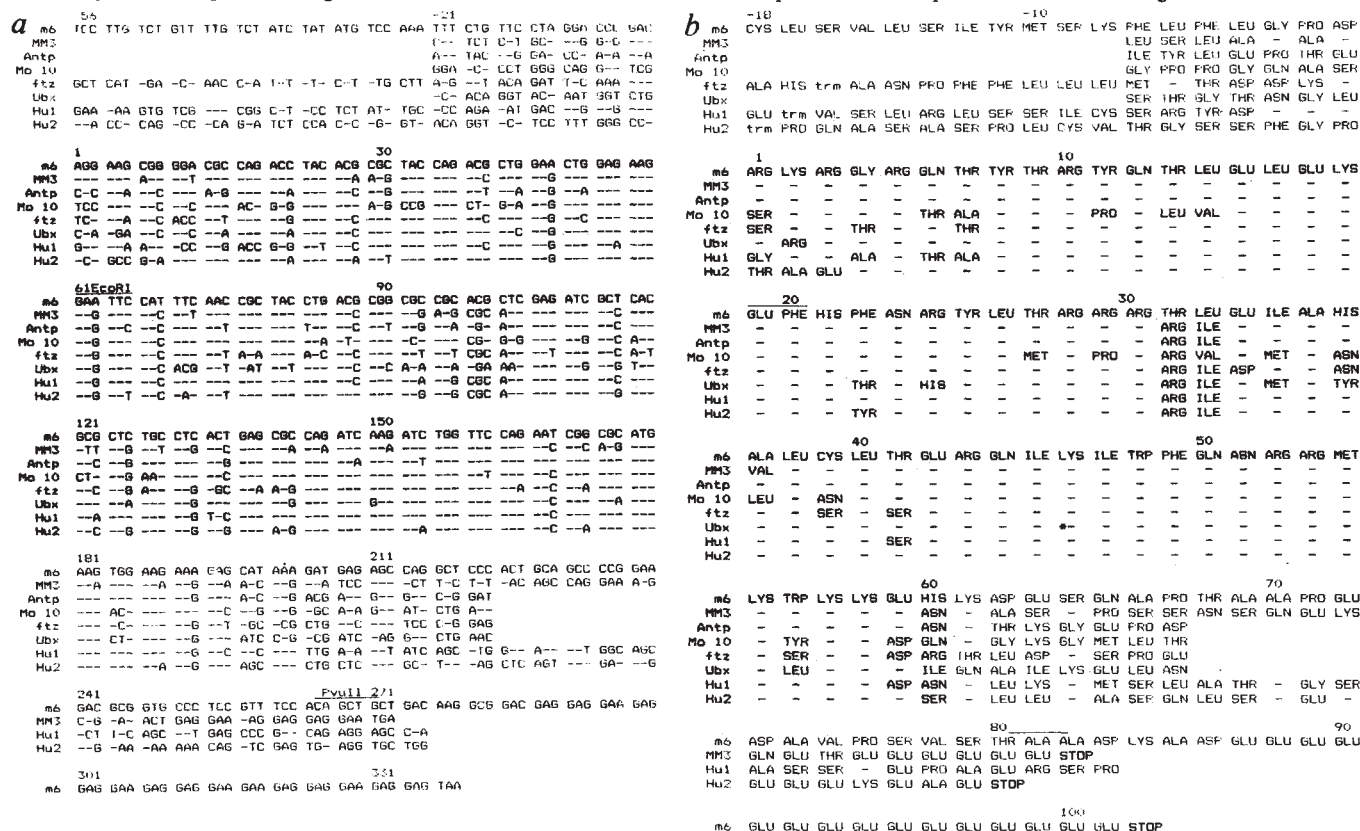


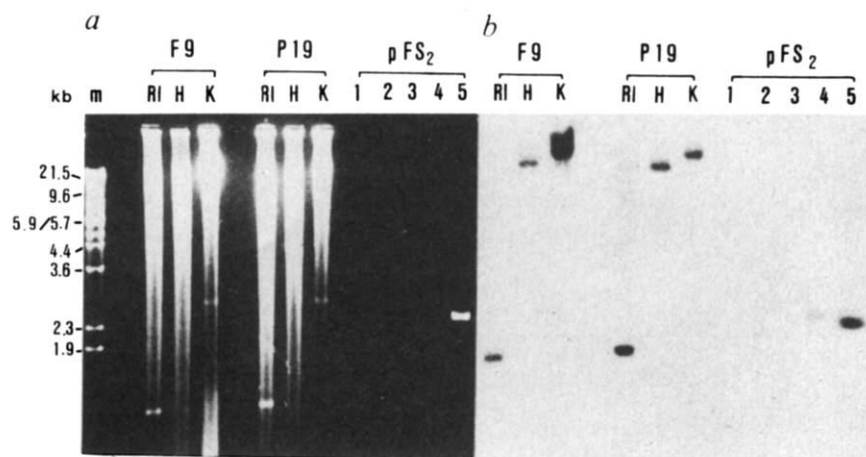
Fig. 3 a, DNA sequence of m6 aligned 5' to 3' with homologous homoeo box regions from mouse (Mo-10)²³, *Xenopus* (MM3)²⁵, *Drosophila* (*Antp*, *ftz*, *Ubx*)¹⁹ and human (*Hu-1* and *Hu-2*)²² DNA. Only those nucleotides that differ from m6 are shown. Conserved nucleotides are indicated by -. b, Conceptual translation of the conserved homoeo box open reading frame in m6 and its comparison with other homoeo box amino-acid domains. Comparison is with amino acids -18 to 83 as numbered in refs 19, 22, 23 and 25. Only non-homologous amino acids are shown directly; conserved amino acids are indicated by -; the asterisk indicates a change in the published amino-acid sequence of *Ubx* (W. Gehring, personal communication).

relative to the amount of β -actin mRNA in each sample as evidenced by hybridization of the blot with a β -actin probe.

To look for RNA encoding potential regulatory proteins induced during differentiation, F9 cells were differentiated by treatment with 5×10^{-7} M retinoic acid and 10^{-3} M dibutyl cAMP. Poly(A)⁺ RNA was obtained at different times during

the first 24 h of differentiation and after 5 days of treatment (Fig. 5b). At 17 h, at least one new RNA (2.4 kb) appeared, although longer exposures of the blot show the presence of an additional low-level 3.2-kb transcript that shows a marginal induction on differentiation of the F9 cells (data not shown). The other two transcripts present in undifferentiated F9 cells, the 1.5- and 2.9-kb mRNAs, remain present during the differenti-

Fig. 4 Hybridization of murine DNA containing a homoeo box (m6-12) to mouse genomic DNA from F9 and P19 cells. **a**, Ethidium bromide staining of cellular DNA. Cell DNA was isolated by treatment of cells with lysis buffer (50 mM Tris-HCl, 20 mM EDTA, 100 mM NaCl, pH 7.5) and digestion with proteinase K (50 μ g ml⁻¹) for 16–24 h. DNAs were purified by phenol extraction and ethanol precipitation. DNAs were cleaved with *Eco*RI (RI), *Hind*III (H) or *Kpn*I (K) and separated on a 0.7% agarose gel (**b**). A Southern blot of DNA in gel **a** was hybridized with nick-translated (m6-12) DNA (1.9 kb *Eco*RI fragment). Various concentrations of uncleaved pFS2 DNA (1, 0.01 ng; 2, 0.1 ng; 3, 1 ng; 4, 10 ng; 5, 100 ng) were included. Hybridization, under stringent conditions, used the same buffer as in ref. 19, with 50% formamide. The blots were washed at least four times in 0.1 $\times$ SSC and 0.1% SDS at 52 $^{\circ}$ C for 15 min each.



ation of the F9 cells, although one of these (2.9 kb) shows a slight increase ($\sim$ twofold) during this period.

To determine whether these mRNAs were present 5 days after induction of differentiation, poly(A⁺) RNA from F9 cells treated with retinoic acid and dibutyl cAMP for short (<24 h) and long (5 days) periods were compared. All the transcripts that hybridized with m6-12 at 17 and 24 h (3.2, 2.9, 2.4 and 1.5 kb) were also present at 5 days, although in apparently reduced quantities (1.5–6 times less) compared with the earlier times. Each value was once again standardized by comparison with the amount of poly(A⁺) RNA blotted for each sample, as determined by β -actin hybridization.

The presence of either homoeo box or flanking sequences in each of these predominant RNA species was examined by hybridization with subfragments contained in the 1.9-kb *Eco*RI fragment, m6-12, which had been used in the initial Northern analyses. Blots of poly(A⁺) RNAs from stem and retinoic acid- and dibutyl cAMP-differentiated (24 h) F9 cells were hybridized with either the *Eco*RI/*Pvu*II 1.7-kb fragment containing flanking sequences (Fig. 6a) or the 0.206-kb *Pvu*II/*Eco*RI fragment that largely contains the murine homoeo box sequences (Fig. 6b). The flanking sequences hybridize with all of the transcripts previously seen in F9 cells at 17 h, 24 h and 5 days after treatment with retinoic acid and dibutyl cAMP. The probe containing the homoeo box, however, hybridizes strongly to only one of these mRNAs, the 2.4-kb mRNA, which is induced shortly after differentiation. In addition, there is a weak hybridization of this probe to the most abundant of these mRNAs, the 1.5-kb species. To determine whether the observed transcripts are encoded by the same or the opposite strand of m6-12 DNA, single-stranded probes were used (Fig. 6c). Hybridization of the noncoding (panel 1) and coding (panel 2) strands showed that all the RNAs detected by the m6-12 (panel 3) probe are complementary to the coding strand and are thus transcribed from the strand encoding the conceptually translated homoeo box open reading frame.

Conclusions

One of the central aims of developmental biology is to understand the molecular details of the mechanisms governing the precise control of differentiation processes. Although little is known about the regulatory mechanisms of mammalian development, the wealth of genetic information³⁴ and the availability of EC stem-cell lines led us to choose these as model systems for the study of mammalian development. Differentiation of F9 stem cells is possible either by treatment with retinoic acid or by transfection with an oncogene (*c-fos*). It remains unclear, however, whether this results from a direct or indirect action of the inducer.

Discovery of common homoeo box sequences in many of the developmental genes in *Drosophila* and developmentally regu-

lated genes of *Xenopus* led us to isolate analogous genes and to study their products in murine EC cells. So far, we have isolated seven mouse phage recombinants that contain homology to the *Drosophila* homoeo boxes and show different restriction patterns. Detailed examination of one λ m clone (λ m6) indicated the presence of at least two homoeo box regions.

Analysis of the nucleotide sequence of one of these regions revealed an extremely well-conserved sequence of 180 bp. Approximately 84% homology to *Antp* can be detected at the nucleotide sequence level. The amino-acid sequences of these conceptually translated proteins are even more highly conserved (95% homology). Amino-acid homology with proteins of known structure (*cro* and *cI*) suggests that the most conserved region includes an α -helical domain (ref. 20 and W. Gehring, personal communication). Interestingly, the most homology was to the homoeo boxes of the *Antp* gene of *Drosophila*, *Xenopus* (MM3) and the human genome, whereas the least homology was to the murine Mo-10 gene. Whether this structural resemblance to the *Antp* gene, MM3 and the human genes also reflects a functional resemblance remains to be determined. Clearly, the m6 homoeo box region is different from the Mo-10 gene. Further comparison of MM3 with m6 reveals conservation of both the homoeo box and a stretch of glutamic acid residues, which lies outside the homoeo box and at the carboxy-terminus of the putative protein. Although the functional significance of this acidic domain remains necessarily speculative, such domains exist in the non-histone nucleosomal proteins HMG1 and HMG2 (ref. 35) and in a neurofilament protein, NF-M (ref. 36).

Before using m6-12 sequences as a probe for Northern analyses, they were characterized by hybridization to genomic blots of murine DNA. Although the m6-12 DNA contains homoeo box sequences present in several copies in the murine genome, Southern analysis showed that it hybridized to unique DNA. The inability of this probe, which primarily contains flanking sequences, to detect other homoeo boxes in the murine genome, may have resulted from the very stringent conditions used in the Southern hybridizations in contrast to the non-stringent conditions of the library screening. Furthermore, possible differences in the sequences of murine homoeo boxes (compare m6 and Mo-10 sequence homology) and the specific activity of homoeo box sequences contained in the probe (<7%) could have precluded their detection.

The probe containing murine homoeo box sequences was used for studies of RNA expression in EK and various stem cells. Surprisingly, F9 stem cells show a strong expression of two differently sized classes of RNA. The same RNA classes can also be seen in EK, PC13 and P19 cells, although in reduced amounts. One interpretation of these results is that F9 cells are already pre-determined and do not represent a true stem cell any longer. Similar conclusions have been drawn by Adamson and Hogan³⁷ on the basis of low-level synthesis of transferrin

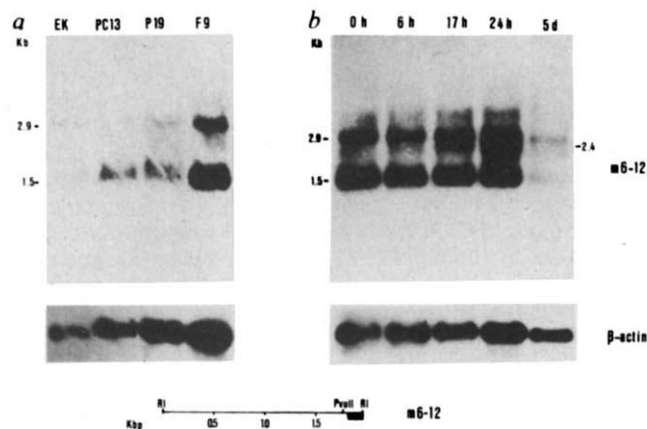


Fig. 5 Hybridization of m6-12 DNA with mRNA from various teratocarcinoma stem cells and differentiated F9 cells. *a*, Poly(A⁺) RNAs from EK, PC13, P19 and F9 cells were hybridized with ³²P-labelled pm6-12 DNA. *b*, RNA obtained from F9 cells treated with retinoic acid (5×10^{-7} M) and dibutyryl cAMP (10^{-3} M) for 6, 17, 24 h and every other day for 5 days were hybridized with ³²P-labelled pm6-12 DNA. Autoradiograms of control hybridizations of each Northern blot (*a* and *b*) with ³²P-β-actin DNA are shown below each corresponding panel. Differences in the intensities of the bands were quantitated by densitometer scanning. The structure of the m6-12 insert in the probe is shown at the bottom of the figure; the solid box represents the homoeo box sequence.

Methods. RNA samples were isolated using the guanidinium thiocyanate method of Chirgwin *et al.*⁴¹ and purified by centrifugation through a 5.7 M CsCl, 25 mM sodium acetate, pH 5.0 cushion. Poly(A⁺) RNAs were obtained from each group by retention on oligo(dT) columns, then 5 μg of poly(A⁺) RNAs was placed in each well and electrophoresed through formaldehyde-agarose gels using 3.7% formaldehyde in MOPS buffer (20 mM morpholinopropane sulphonic acid, 50 mM sodium acetate, 10 mM EDTA pH 7.0). RNAs were blotted onto Gene Screen Plus and hybridized to nick-translated plasmid containing m6-12 DNA (lower portion of the figure) and then to β-actin DNA. Briefly, Northern blots were hybridized with ³²P-DNA in 50% formamide, 1.0 M NaCl, 1% SDS and 100 μg ml⁻¹ denatured salmon sperm DNA. Blots were washed twice in 2×SSC at room temperature, twice in 2×SSC, 1% SDS at 60°C, and twice in 0.1×SSC at room temperature.

and expression of epidermal growth factor receptors in F9 stem cells, markers which are not usually present on other stem cells.

We were also interested to determine whether differentiation of F9 cells is accompanied by the differential expression of murine genes. RNA produced by F9 cells at various times after differentiation into parietal endoderm showed the early appearance of an additional transcript (2.4 kb) that hybridized with both the homoeo box and flanking sequences contained in the m6-12 probe. The transcripts present in F9 stem cells and those which increase so rapidly during the first 24 h of differentiation persist well into the time when F9 cells express markers of differentiated cells (5 days), although in decreased quantities compared with another transcript such as β-actin. RNA from macrophage, a terminally differentiated cell type, did not accumulate detectable amounts of these mRNAs (data not shown). To establish the relationship between the detected RNAs, the 206-bp *PvuII*/*EcoRI* fragment containing the murine homoeo box and an additional 90 bp was hybridized with mRNAs from F9 stem and differentiated (24 h) cells and the 2.4-kb mRNA was detected, suggesting that the transcript contains homoeo box sequences. Although there is also a weak hybridization to the most abundant mRNA (1.5 kb; data not shown), the intensity of this hybridization, compared with that seen with the 2.4-kb mRNA, suggests that not all of the probe sequences are present in this transcript. This hybridization could result either from homology with the 206-bp probe or from a minor contamination of this probe with the larger 1.7-kb

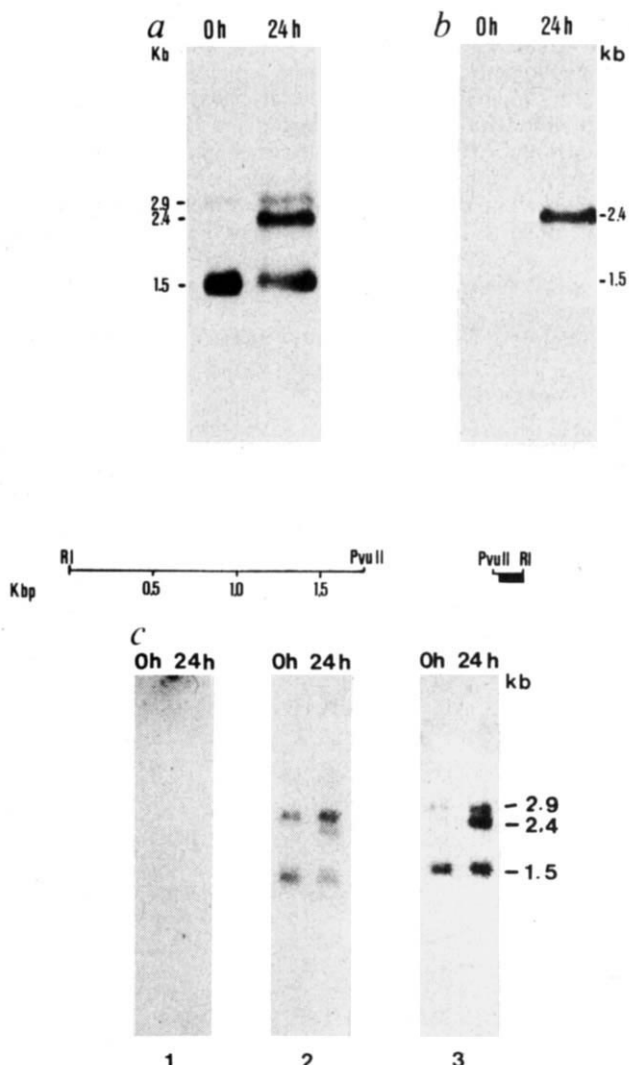


Fig. 6 Hybridization of the flanking (*a*) and homoeo box (*b*) sequences contained in m6-12 DNA and (*c*) single-stranded M13 m6-12 DNA probes to stem-cell (0 h) and differentiated F9 (24 h) cellular RNA. The structures of the probes are shown at the bottom of the figure (solid box represents homoeo box sequences).

Methods. Poly(A⁺) RNAs from F9 stem cells and those treated with retinoic acid and dibutyryl cAMP for 24 h were obtained as described in Fig. 5 legend. Samples (5 μg) were separated by electrophoresis in formaldehyde-agarose gels, blotted onto Gene Screen Plus and hybridized to either the nick-translated 1.7-kb *EcoRI*/*PvuII* fragment (*a*), the 206-bp *PvuII*/*EcoRI* fragment (*b*) from m6-12 or to the noncoding (1) or coding (2) strands of M13mp10 or mp11 recombinants (*c*) containing the 1.3-kb *SacI*/*EcoRI* insert from m6-12. As a control, nick-translated pm6-12 (3) was also hybridized to F9 stem-cell (0 h) and differentiated F9 cell (24 h) RNA (*c*). The M13 single-stranded DNA probes were labelled by incorporation of ³²P-dCTP using primer extension, denatured in 0.1 M NaOH and electrophoresed in alkali agarose gels. Labelled single strands were recovered by electroelution.

PvuII/*EcoRI* fragment. Because the larger *PvuII*/*EcoRI* fragment containing the remaining m6-12 sequences hybridizes with all of the transcripts previously seen following differentiation, the 3.2-, 2.9-, 2.4- and 1.5-kb mRNAs presumably all contain these flanking sequences. It is likely, however, that although these mRNAs are transcribed from the strand of DNA that encodes the homoeo box protein open reading frame and map closely on the mouse genome, only the 2.4-kb mRNA contains homoeo box sequences. All of the transcripts that hybridize to the pm6-12 probe were still detected after washing the Northern blots under very stringent conditions (0.1×SSC at 60°C, data not shown).

In attempts to determine whether the induced 2.4-kb mRNA is specific for the production of parietal endoderm, we have performed experiments using several cell types: F9 cells differentiated by treatment with retinoic acid and aggregation into visceral endoderm, P19 cells differentiated into myoblasts^{38,39} and mouse embryos isolated at different stages of development. Experiments with retinoic acid-treated F9 cells (visceral endoderm) showed a 2.4-kb mRNA on differentiation similar to that present in parietal endoderm. P19 cells differentiated into myoblasts, as well as mouse embryonic and extra-embryonic tissues, also express mRNAs that hybridize to m6-12 sequences.

We intend to use the different EC cells as model systems to

study the role of various homoeotic genes in the differentiation of these cells.

We thank W. McGinnis and W. Gehring for the homoeo box plasmids, H. Lehrach for the murine library, B. Paterson for the β -actin plasmid, C. Stewart and E. Wagner for EK poly(A⁺) RNA and critical reading of the manuscript and M. Kessel and W. Franke for helpful discussions. We also thank E. Vossbans-Bosbach for technical assistance and R. Franklin for typing the manuscript. A.M.C.-P. holds a fellowship from the Alexander von Humboldt Stiftung; S.D.V. a Fulbright Fellowship and K.C. was supported by the Bundesministerium für Forschung und Technologie (BCT0364/1). This research was supported by the Deutsche Forschungsgemeinschaft (DFG Ba 384/18-4).

Received 2 January; accepted 11 March 1985.

- Kleinsmith, L. J. & Pierce, G. B. Jr *Cancer Res.* **24**, 1544-1551 (1964).
- Martin, G. R. *Science* **209**, 768-776 (1980).
- Bernstine, E. G., Hooper, M. L., Grandchamp, S. & Ephrussi, B. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3899-3903 (1973).
- Solter, D. & Knowles, B. B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5565-5569 (1978).
- Artzt, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2988-2992 (1973).
- Stern, P. L. et al. *Cell* **14**, 775-783 (1978).
- Paulin, D., Jakob, H., Jacob, F., Weber, K. & Osborn, M. *Differentiation* **22**, 90-99 (1982).
- Sherman, M. I. & Miller, R. A. *Dev. Biol.* **63**, 27-34 (1978).
- Strickland, S. & Mahdavi, V. *Cell* **15**, 393-403 (1978).
- Adamson, E. D., Gaunt, S. J. & Graham, C. F. *Cell* **17**, 469-476 (1979).
- Hogan, B. L. M., Taylor, A. & Adamson, E. *Nature* **291**, 235-237 (1981).
- Müller, R. & Wagner, E. F. *Nature* **311**, 438-442 (1984).
- Kahan, B. & Adamson, E. D. *Cold Spring Harb. Conf. Cell Prolif.* **10**, 131-141 (1983).
- Boller, K. & Kemler, R. *Cold Spring Harb. Conf. Cell Prolif.* **10**, 39-49 (1983).
- Strickland, S., Smith, K. K. & Marotti, K. R. *Cell* **21**, 347-355 (1980).
- Levine, R. A., La Rosa, G. J. & Gudas, L. J. *Molec. cell. Biol.* **4**, 2142-2150 (1984).
- Stacey, A. J. & Evans, M. J. *EMBO J.* **3**, 2279-2285 (1984).
- McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. & Gehring, W. J. *Nature* **308**, 428-433 (1984).
- McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. *Cell* **37**, 403-408 (1984).
- Laughon, A. & Scott, M. P. *Nature* **310**, 25-31 (1984).

- Scott, M. P. et al. *Cell* **35**, 763-776 (1983).
- Levine, M., Rubin, G. M. & Tjian, R. *Cell* **38**, 667-673 (1984).
- McGinnis, W., Hart, C. P., Gehring, W. J. & Ruddle, F. H. *Cell* **38**, 675-680 (1984).
- Carrasco, A. E., McGinnis, W., Gehring, W. J. & De Robertis, E. M. *Cell* **37**, 409-414 (1984).
- Müller, M. M., Carrasco, A. E. & De Robertis, E. M. *Cell* **39**, 157-162 (1984).
- Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. J. *molec. Biol.* **170**, 827-842 (1983).
- Garber, R. L., Kuroiwa, A. & Gehring, W. J. *EMBO J.* **2**, 2027-2036 (1983).
- Kuroiwa, A., Hafen, E. & Gehring, W. J. *Cell* **37**, 825-831 (1984).
- Southern, E. J. *molec. Biol.* **98**, 503-517 (1975).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- McBurney, M. W. & Rogers, B. J. *Dev. Biol.* **89**, 503-508 (1982).
- Evans, M. J. & Kaufman, M. H. *Nature* **292**, 154-156 (1981).
- Gluecksohn-Waelsch, S. *Cold Spring Harb. Conf. Cell Prolif.* **10**, 3-13 (1983).
- Walker, J. M., Gooderham, K., Hastings, J. R. B., Mayes, E. & Johns, E. W. *FEBS Lett.* **122**, 264-270 (1980).
- Geisler, N., Fischer, S., Vandekerckhove, J., Plessmann, U. & Weber, K. *EMBO J.* **3**, 2701-2706 (1984).
- Adamson, E. D. & Hogan, B. L. M. *Differentiation* **27**, 152-157 (1984).
- Edwards, M. K. S., Harris, J. F. & McBurney, M. W. *Molec. cell. Biol.* **3**, 2280-2286 (1983).
- Dony, C., Kessel, M. & Gruss, P. *Cell Differentiation* (in the press).
- Umene, K. & Enquist, L. W. *Gene* **13**, 251-268 (1981).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).

LETTERS TO NATURE

Are galaxies more strongly correlated than clusters?

Alexander S. Szalay* & David N. Schramm

The University of Chicago, Chicago, Illinois 60637, USA and Fermilab, Batavia, Illinois 60510, USA

One of the most powerful tools used in attempts to understand the structure of the Universe is the correlation function $\xi(r)$, the excess probability over random that there are two objects separated by a distance r . In particular, distributions of galaxies and of clusters of galaxies have been investigated using this parameter. Here we show that if the amplitudes of the cluster-cluster correlation function is made dimensionless, systematic changes with cluster richness vanish, implying scale invariance in the clustering process. The dimensionless galaxy-galaxy correlation seems stronger, implying gravitational enhancement on smaller scales.

It has been established¹ that for distances out to 10 Mpc, the galaxy distribution has a power-law correlation of the form $\xi_{gg}(r) = \alpha_{gg}(hr)^{-1.8}$, where h is the Hubble constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Recently, the spatial correlation function of Abell clusters has been determined directly^{2,3}, following earlier work on the angular correlations⁴. It was noted that rich ($R \geq 1$) clusters of galaxies correlate with the same power law as galaxies, but with a significantly greater amplitude, $\alpha_{cc} = 18\alpha_{gg}$, which is even higher for richer ($R \geq 2$) clusters. Clusters poorer than most Abell clusters were identified in the Lick catalogue by a numerical algorithm⁵. The correlations had the same slope and the amplitude was between that of galaxies and that of the Abell clusters, in agreement with the richness trend (see Table 1).

An attempt to explain the behaviour of α_{cc} has been made by Kaiser⁶ applying the statistics of rare events. If the density perturbations are described by a random gaussian field and the regions where clusters form correspond to densities higher than a certain threshold, the correlation function of the points above the 'clipping level' is a constant factor times the correlation function of all the points. By filtering out from the initial power spectrum the scales smaller than clusters and choosing the appropriate threshold, it was possible to match both the number density and enhanced correlations of the Abell clusters. To explain the factor of 20 in α_{cc} the clusters have to be $\geq 3\sigma$ points. In this picture $\xi_{cc}(r)$ will have the same functional form as the galaxy autocorrelations. The model may have some difficulties here, because ξ_{gg} seems to be negative at 20-Mpc radii, but ξ_{cc} is positive out to at least 100 Mpc.

The r.m.s. peculiar velocities and the correlation amplitude of the mass are related; both originate from the power spectrum of perturbations. The clusters cannot represent the true mass autocorrelations in the universe, because then we would expect to see enormous (a few thousand kilometres per second) peculiar velocities for the clusters, which does not seem to be the case.

Whenever similar systems of different scales are compared, it is better to use dimensionless quantities. The correlation amplitude α_{cc} is dimensional; it is $\xi_{cc}(r)$ at a distance $hr = 1$ Mpc. It can be made dimensionless by using an intrinsic unit of length for each sample of different richness. The natural length that appears in these point catalogues is the mean separation L , determined by the density $n = L^{-3}$. This mean separation is indeed intrinsic because there is a unique one-to-one correspondence between richness and the value of L , which is because the richer samples are always subsets of the poorer catalogue. Therefore, labelling the different richness selections by L , the density of each sample is one, and the value of the correlation function at L (expressing distance in units of L) is also dimensionless: $\beta(L) = \xi(L) = \alpha(L)L^{-1.8}$.

* Permanent address: Eotvos University, Budapest, Hungary.

Table 1 The dimensional correlation amplitude $\alpha(L)$, compared with the dimensionless $\beta(L) = \alpha(L)L^{-1.8}$ for groups and clusters of increasing richness

Sample	$\alpha(L)$	n	L	$\beta(L)$
Galaxies	20	8.0×10^{-3}	5	1.10
Schechterman	180	3.6×10^{-5}	30	0.39
BS $R \geq 1$	360	6.0×10^{-6}	55	0.27
BS $R \geq 2$	1,080	2.0×10^{-6}	80	0.41

n , The comoving spatial density in Mpc^{-3} ; $L = n^{-1/3}$, the mean separation in Mpc. The spatial density of galaxies is given in ref. 15.

The constant slope in the power law behaviour of $\xi(r)$ indicates scale invariance of the process responsible for galaxy formation. In the case of galaxies, the typical distances are a few Mpc, whereas for clusters the same law holds to distances > 50 Mpc. If galaxy formation were completely scale invariant, one would expect that all dimensionless correlation amplitudes would be equal.

Figure 1 is a plot of the dimensionless amplitude $\beta(L)$ for all the samples, including the different richness classes of Bahcall and Soneira². Note that $\beta(L)$ for galaxies is 1.1, a factor of three higher than for the cluster samples, which are all consistent with 0.35. As far as clusters are concerned, the scale invariance seems to hold. The correlation amplitudes of the cluster catalogues are fairly well known, but their densities are less certain. Because L and thus $\beta(L)$ depend on the density, we plotted conservative error limits on Fig. 1 corresponding to densities 1.5 times above and below the quoted value. Thus compared, galaxies are more strongly correlated than clusters, the opposite conclusion to that obtained by comparing $\alpha(L)$.

The relative constancy of $\beta(L)$ can be understood in general terms as a signature of scale invariance. If there is a nonlinear process (besides gravity) modulating galaxy formation and this process is scale invariant, the created pattern will have a single power law correlation function⁷, $\xi(r) = \beta(r/L)^{D-3}$, the slope of which would be related to the geometry of the pattern, namely to its fractional (or 'fractal') dimension D . This $\xi(r)$ will have the correct scaling property. Thus, to explain the observed slope, a fractal dimension of $D = 1.2$ is required, but we do not know yet what physical process can generate the correct D . (We discuss several possibilities below.)

Small-scale gravitational clustering may break the scale invariance and increase the dimensionless correlation amplitude for galaxies. In one extreme case, we can approximate this process by assuming that the pattern imposed on the galaxy distribution is uncorrelated with the gravitational motion. Then the observed correlations are a superposition of the gravitational and fractal correlations, with $\xi_{\text{frac}}(r) = 0.35(r/L)^{-1.8}$, which is scale invariant, and $\xi_{\text{grav}}(r)$ a flat function of r out to ~ 10 Mpc, independent of L . ξ_{grav} is the Fourier transform of the power spectrum: $1 + \xi_{\text{obs}}(r) = [1 + \xi_{\text{grav}}(r)][1 + \xi_{\text{frac}}(r)]$. If $\xi_{\text{grav}}(0) \approx 3$, then we approximately reproduce the observed factor of three in $\beta(L)$.

In the other extreme the process is the nonlinear 'clipping' of the gravitationally induced overdensities⁸, in which case the 'fractal' and the gravitational perturbations are fully correlated. Kaiser has also found a limited scale invariance for that particular model. If the selection is simply the result of a threshold in the initial perturbations, which are otherwise scale invariant, the resulting regions, the so-called 'excursion sets', will have a scale-invariant geometry and will be fractals⁸. If the universe is dominated by 'cold' dark matter, the fluctuation spectrum has an asymptotic form of k^{-3} , where k is the wavenumber. The larger galaxy correlations in this case can be attributed to effects of nonlinear gravitational clustering on small scales. From comparisons of the low r.m.s. galaxy velocities and the high correlation function of galaxies, it seems that such a 'bias' is useful⁹.

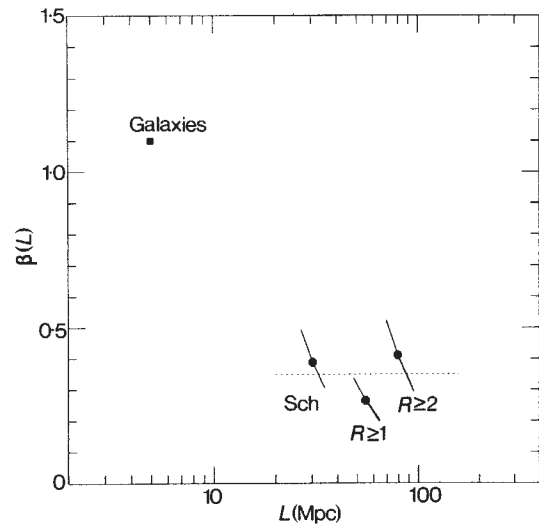


Fig. 1 The dimensionless correlation amplitude $\beta(L)$ is shown as a function of the mean separation L . Note that all three cluster catalogues are consistent with 0.35 (dashed line), whereas the galaxy amplitude is 1.1. The error-bars in the cluster data represent a factor of 1.5 up/down uncertainty in the density.

All these mechanisms have the advantage of increasing the galaxy correlations while leaving the velocities intact.

Schulman and Seiden¹⁰ have already suggested similarities between condensed matter physics and galaxy formation. Their conclusion, from simple analogies with short range interaction spin systems, was that only fractals with $D \sim 2$ can be obtained.

Gas dynamical processes are bound to occur at some level during galaxy formation. The formation and explosion of first generation Pop III stars^{11,12} could have major consequences for the entire galaxy formation scenario. Although the energies available in each explosion are insufficiently large, percolation may occur, creating random structures of 'infinite' length that can produce the required behaviour.

Another possibility is relativistic strings^{13,14}, which have attracted much attention recently. If present, they may create patterns with a fractal dimension close to unity.

In conclusion, we propose that galaxy and cluster correlation functions be compared in dimensionless units. The dynamic range of correlation amplitudes has been reduced from 50 to 3. All cluster correlations are consistent with the same scale-invariant amplitude. That in those units galaxies are more strongly clustered may be because of small-scale gravitational clustering, whereas the slope and the dimensionless amplitude may reflect a scale invariant process responsible for creating luminous galaxies in certain regions of the universe.

We thank N. Bahcall, N. Kaiser and T. Vicsek for useful discussions. This work was supported in part by the NSF, the National Aeronautics and Space Administration and the US Department of Energy.

Received 7 February; accepted 28 February 1985.

1. Groth, E. J. & Peebles, P. J. E. *Astrophys. J.* **217**, 385-405 (1977).
2. Bahcall, N. A. & Soneira, R. A. *Astrophys. J.* **270**, 20-38 (1983).
3. Klypin, A. A. & Kopylov, A. I. *Sov. Astr. Lett.* **9**, 41-46 (1983).
4. Hauser, M. G. & Peebles, P. J. E. *Astrophys. J.* **185**, 757-786 (1973).
5. Schechterman, S. A. *Astrophys. J.* (in the press).
6. Kaiser, N. *Astrophys. J.* **284**, L9-L12 (1984).
7. Mandelbrot, B. B. *Fractals* (Freeman, San Francisco, 1977).
8. Adler, R. J. *The Geometry of Random Fields* (Wiley, New York, 1981).
9. White, S. D. M. *Proc. Inner Space/Outer Space* (University of Chicago Press, 1985).
10. Schulman, L. S. & Seiden, P. E. *IBM Rep. RC-10369* (1984).
11. Carr, B. J., Bond, J. R. & Arnett, D. *Astrophys. J.* **277**, 445-469 (1984).
12. Carr, B. J. & Ikeuchi, S. *Kyoto Preprint RIFP-550* (1984).
13. Turok, N. & Albrecht, A. *ITP Santa Barbara Preprint* (1984).
14. Turok, N. & Schramm, D. N. *Nature* **312**, 598-599 (1984).
15. Davis, M. & Huchra, J. *Astrophys. J.* **254**, 437-454 (1982).

A new non-thermal galactic radio source with a possible binary system

E. Fürst*, W. Reich*, P. Reich*, Y. Sofue† & T. Handa†

* Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-5300 Bonn 1, FRG

† Nobeyama Radio Observatory, Tokyo Astronomical Observatory, University of Tokyo, Minamisaku, Nagano 384-13, Japan

Helfand and Becker¹ recently discussed the unusual filamentary appearance of the galactic radio sources G357.7-0.1 and G5.3-1.0, both of which had originally been classified as supernova remnants (SNRs)^{2,3}, and proposed that these sources belong to a new class of non-thermal radio sources that originate in accreting binary systems containing a neutron star or a black hole. Becker and Helfand⁴ have pointed out that other galactic sources may be identified as members of this class; we report here our claim that a new non-thermal galactic object (G18.95-1.1), detected in the recently published 2.695-GHz galactic plane survey⁵, is a possible candidate. That the integrated flux density spectral index $\alpha = -0.4$ ($S_\nu \sim \nu^\alpha$) and the polarization is $\approx 2.5\%$ at 4.75 GHz, proves the non-thermal nature of the new source; morphologically, a classification of this object as a SNR seems impossible. G18.95-1.1 consists of various arcs all pointing towards a central radio peak; we suggest that the object is a binary system containing a compact component, located at or close to this radio peak, accelerating electrons to relativistic energies—observed in the arcs by the synchrotron emission of the accelerating electrons.

The source G18.95-1.1 has subsequently been observed at 1.42 GHz and 4.75 GHz with the Effelsberg 100-m telescope and at 10 GHz with the Nobeyama 45-m dish. The 4.75-GHz map is shown in Fig. 1 and the spectrum is plotted in Fig. 2. Both the spectral index and the polarization reveal the non-thermal nature of G18.95-1.1.

How do we classify this object? The peak radio intensity is close to the centre of the object and shows a bar-like structure with peaks near each end of the bar (Fig. 1). Even the highest available angular resolution does not allow separation of any compact source from the diffuse emission. The grey-scale radiograph in Fig. 3 allows the determination of the source structure

in much greater detail. From this presentation it is apparent that G18.95-1.1 consists of various arcs pointing towards the radio peak slightly off-centre to the north-east. Clearly, the morphology of G18.95-1.1 defies any classification as a SNR and also contradicts the general appearance of radio galaxies. Its arcs may be of similar origin to the filaments in G357.7-0.1 discussed by Helfand and Becker¹. We therefore suggest a binary system containing a compact component as the source of G18.95-1.1. This possible centre of activity should be expected to be located in the central bar, perhaps at or close to the peak emission north-east of the centre. If the central location of this source is not the result of projectional effects, the velocity of the binary system relative to the filamentary structure is low, or the source is much younger than G357.7-0.1 and G5.3-1.0.

The estimation of source parameters (for example, age, distance) is hindered by the insufficient data available. We suggest that deep optical observations (no corresponding emission could be found in the Palomar plates) and a search for X-ray emission be made. At radio wavelengths, more sensitive polarization

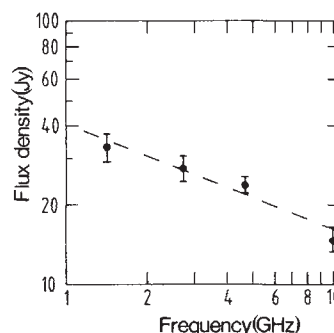


Fig. 2 The radio spectrum of G18.95-1.1. The integrated flux densities are: $S(1.42 \text{ GHz}) = 32.9 \text{ Jy}$, $S(2.695 \text{ GHz}) = 27.4 \text{ Jy}$, $S(4.75 \text{ GHz}) = 23.8 \text{ Jy}$ and $S(10 \text{ GHz}) = 14.6 \text{ Jy}$; the typical error is 10%. The mean flux density spectral index $\alpha = -0.4$. There may be a steepening of the spectrum from $\alpha = -0.4$, in the inner part near the radio peak, to $\alpha = -0.5$ at the outer arcs, but the accuracy of the spectral index is 0.2. The errors are mainly attributable to the uncertainty in the estimation of the base levels.

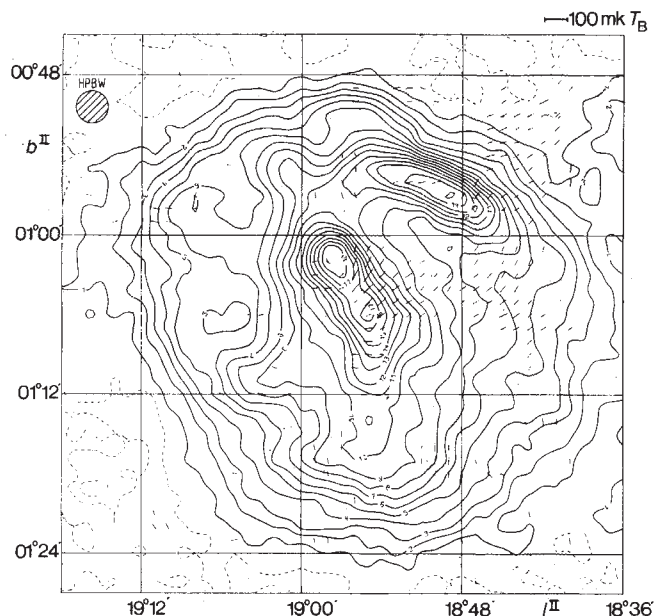


Fig. 1 4.75-GHz map of G18.95-1.1. Contour steps are 50 mK brightness temperature T_B . The angular resolution is 2.4 arc min. Superposed are the linear polarization intensity E-vectors. The integrated polarization percentage is 2.5%.



Fig. 3 The grey-scale map of G18.95-1.1.

measurements together with improved angular resolution (possibly using the Very Large Array) and a search for a pulsar will certainly help to understand the nature of G18.95-1.1. If our assumption of a binary system as the source of G18.95-1.1 is confirmed, this object will provide important information on the interaction of binary systems with the surrounding interstellar medium.

We thank Professors R. Wielebinski and W. Kundt and Dr W. Sieber for helpful comments.

Received 25 January; accepted 28 February 1985.

1. Helfand, D. J. & Becker, R. H. *Nature* **313**, 118-119 (1985).
2. Milne, D. E. *Aust. J. Phys.* **32**, 83-92 (1979).
3. Green, D. A. *Mon. Not. R. astr. Soc.* **209**, 449-478 (1984).
4. Becker, R. H. & Helfand, D. J. *Nature* **313**, 115-118 (1985).
5. Reich, W., Fürst, E., Steffen, P., Reif, K. & Haslam, C. G. T. *Astr. Astrophys. Suppl.* **58**, 197-248 (1984).

A higher limit to the mean galaxy mass allowed by gravitational lens image distortion

Steven Phillipps

Department of Applied Mathematics and Astronomy,
University College, PO Box 78, Cardiff CF1 1XL, UK

Tyson *et al.*¹, in a recent discussion of the gravitational lens distortion, by foreground galaxies, of the images of distant ('background') galaxies, claim that their non-detection of this effect implies a low mean mass for the foreground galaxies. The effect they sought is a change in the coherent ellipticity, that is, the ratio of the angular diameters orthogonal and parallel to the separation vector to the foreground galaxy. Using samples of galaxies in different ranges of magnitude to represent the 'foreground' ($19 < J < 21.5$) and 'background' ($22.5 < J < 23.5$) galaxies, they found no significant effect and therefore deduced a 2σ upper limit to the mean mass of the foreground galaxies of only $2.8 \times 10^{11} H^{-1} M_{\odot}$ out to a radius of $80 H^{-1}$ kpc (where H is Hubble's constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$). The corresponding mass-to-light ratio and the known luminosity density of the Universe then requires that the mass density in the form of galaxies is $\Omega_{\text{gal}} < 0.03$, which is in marked contrast to the values $\Omega_{\text{gal}} \sim 0.2$ obtained by various dynamical arguments, namely, the rotation curves of individual galaxies, the virial equation applied to groups and clusters and the cosmic virial theorem²⁻⁵. I show here that, by including the effect of galaxy clustering, the limit on Ω_{gal} is indeed raised to $\Omega_{\text{gal}} \leq 0.2$, in accordance with the dynamical determinations.

Tyson *et al.*¹ noted that for the two magnitude ranges they used, there should be little overlap in the distances of their two sets of galaxies, and substantiated this by Monte Carlo experiments using a Schechter⁶ luminosity function. They claimed therefore that almost all observed pairs should be a genuine background-foreground pair (the lensing effect is insignificant if the two galaxies are close together). However, this step is invalid if galaxy clustering is taken into account, because for any observed pair, what we really require is the conditional probability of finding a 'background' (that is, faint) galaxy at a particular distance given the presence of the foreground galaxy, since this becomes significantly confused by nearby clustered low-luminosity galaxies. It is easy to see that this has a substantial effect by considering the standard form for the galaxy autocorrelation function⁷, $\xi(s) = (s/r_0)^{-\gamma}$, with index $\gamma = 1.8$ and scale length $r_0 \approx 5 H^{-1} \text{ Mpc}$. Even at the maximum pair separation used by Tyson *et al.*, corresponding to $220 H^{-1} \text{ kpc}$ at the mean distance $D = 720 H^{-1} \text{ Mpc}$ of the foreground galaxies, the clustering increases the number of expected neigh-

bours by factors of several hundred relative to that for a random distribution.

More exactly, we can integrate the correlation function to show that the number of galaxies per unit solid angle at separation θ from a given foreground galaxy, distance D from the observer, will be⁸

$$N' = N + D^2 \Phi(M) G(\gamma) r_0^\gamma (\theta D)^{1-\gamma}$$

where N is the mean surface density, Φ is the integral luminosity function and $G = 3.678$ for $\gamma = 1.8$. Now, from Tyson *et al.*, we find the mean background density to be $5,670$ galaxies per deg^2 or $N = 1.86 \times 10^7 \text{ rad}^{-2}$. Also, with the previously specified values of D , G , r_0 and γ and taking⁹ $\Phi \approx 0.05 H^3 \text{ Mpc}^{-3}$ (in this context Φ is the number density of galaxies in the magnitude range $-16.5 \geq M_B - 5 \log H \geq -17.5$ corresponding to $22.5 < J < 23.5$ at $D = 720 H^{-1} \text{ Mpc}$), the second term on the right-hand side has the value $1.7 \times 10^8 \theta_s^{-0.8}$, where θ_s is the separation in arc seconds.

In other words, the fraction of genuine background-foreground pairs that are able to show a measurable lensing effect is only $1/(1 + 9 \theta_s^{-0.8})$ so that the mean distortion will be diluted by a factor ranging from 4.0 at $\theta = 4 \text{ arc s}$ to 1.5 at $\theta = 40 \text{ arc s}$. (Note that the ratio will be the same even if there is a bias against detecting close pairs, because of the presence of the bright foreground galaxy, thus reducing the apparent clustering.) As the effect sought is greatest at small separations, this obviously greatly limits the power of the test. When the dilution factor is included, even the most massive model considered by Tyson *et al.* (outer radius $190 H^{-1} \text{ kpc}$, equivalent circular velocity 300 km s^{-1} , that is, mass $= 2.1 H^{-1} \times 10^{12} M_{\odot}$) is within 2.5σ of the observed points for all except one of the bins in θ . It may therefore be prudent, at present, to take a conservative upper limit of $\sim 2 \times 10^{12} H^{-1} M_{\odot}$ for the mass of a mean foreground galaxy, which raises the limit on the density parameter to $\Omega_{\text{gal}} \leq 0.2$. We note also that if the clustered galaxies are preferentially aligned¹⁰, then the situation is even worse than we have assumed and almost any density parameter will be allowed, because this effect will work in opposition to the lens effect that causes an orthogonal elongation.

I thank Professors M. J. Disney and B. F. Schutz and Drs E. Balbinski, M. G. Edmunds, A. H. Nelson and A. P. Whitworth for valuable discussions.

Received 23 November 1984; accepted 26 February 1985.

1. Tyson, J. A., Valdes, F., Jarvis, J. F. & Mills, A. P. *Astrophys. J.* **281**, L59-L62 (1984).
2. Rubin, V. C., Ford, W. K. & Thonnard, N. *Astrophys. J.* **225**, L107-L112 (1978).
3. Faber, S. M. & Gallagher, J. S. A. *Rev. Astr. Astrophys.* **17**, 135-187 (1979).
4. Peebles, P. J. E. in *Physical Cosmology* (eds Balian, R., Audouze, J. & Schramm, D. N.) (North-Holland, Amsterdam, 1980).
5. Bean, A. J., Efstathiou, G., Ellis, R. S., Peterson, B. A. & Shanks, T. *Mon. Not. R. astr. Soc.* **205**, 605-624 (1983).
6. Schechter, P. *Astrophys. J.* **203**, 297-306 (1976).
7. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton University Press, 1980).
8. Phillipps, S. *Mon. Not. R. astr. Soc.* **212**, 657-661 (1985).
9. Felton, J. E. *Astr. J.* **82**, 861-878 (1977).
10. Djorgovski, S. *Astrophys. J.* **274**, L7-L12 (1983).

Synthesis of ethylene and ethane by partial oxidation of methane over lithium-doped magnesium oxide

Tomoyasu Ito & Jack H. Lunsford

Department of Chemistry, Texas A&M University, College Station, Texas 77843, USA

The partial oxidation of methane into more useful chemicals such as methanol, ethylene and benzene has been investigated extensively, although yields for these products have been poor¹⁻⁴. Moreover, in several of these processes the required oxidant is N_2O rather than O_2 . Recent work⁵ in our laboratory has demon-

Table 1 Results for the partial oxidation of methane at 720 °C over Li/MgO

Run no.	Reactants		Products				Conversion† %	C ₂ selectivity %	C ₂ yield %
	CH ₄	O ₂	CO ₂	CO	C ₂ H ₄	C ₂ H ₆			
1	59	29	11.2	0.0	3.5	2.2	37.8	50.3	19.0
2	58	36	13.7	0.2	3.7	2.1	42.8	45.4	19.4
3	89	29	11.1	0.0	4.6	3.1	29.1	58.1	16.9
4	218	117	51.2	1.5	15.4	7.4	37.5	46.5	17.4

In earlier runs (not listed), Li/MgO was used at 770 °C. Values for C₂ selectivity and C₂ yield were based on the amount of methane that had reacted. C₂ yield is defined as the product of conversion and C₂ selectivity.

strated that lithium-doped magnesium oxide (Li/MgO) in the presence of O₂ has high activity for abstracting H from CH₄ to form $\cdot\text{CH}_3$ radicals. This suggests that C₂H₆ and C₂H₄ (C₂ compounds) are produced by a coupling between two gaseous $\cdot\text{CH}_3$ radicals formed on this catalyst. We report here our success in converting CH₄ to C₂ compounds in high yields in conventional catalytic conditions.

Powder catalysts of Li/MgO, containing 3 and 7 wt% Li, were prepared from Li₂CO₃ and MgO⁵ and then activated at 465 °C in an oxygen flow. The catalytic experiments were performed at 1 atm in a conventional flow reactor of the type described previously⁴. A reacting gas mixture of CH₄, O₂ and He (a diluent) was introduced at a flow rate of 0.83 ml s⁻¹. All of the gases were analysed by gas chromatography except HCHO, which was analysed by a titration technique⁴.

High yields for the C₂ compounds were attained when the ratio of O₂ to CH₄ in a reactant mixture was ~0.5. The results, which were obtained at 720 °C over 4 g of 7% Li/MgO, are summarized in Table 1. Besides C₂H₄, C₂H₆, CO₂ and CO, CH₃OH and HCHO also were produced, but in only trace amounts. By comparison, the conversion of CH₄ was 0.2% at 720 °C in the same experimental conditions except for the absence of the catalyst. Pure MgO catalyst (1 g), prepared by the same method as Li/MgO but without Li₂CO₃, produced the C₂ compounds with a low selectivity of 6% at a conversion level of 17% at 700 °C.

A yield of 19% for the C₂ compounds was obtained with 50.3% selectivity at 37.8% conversion (run 1), which is considerably better than results reported in the literature for other metal oxide catalysts^{1,3}. Almost equal amounts of CH₄ and O₂ reacted with each other, and most of the O₂ was consumed during the reaction in these experimental conditions. By feeding a reactant gas mixture containing O₂ in a higher ratio than that of run 1,

we could attain a higher conversion of CH₄, but the selectivity for the C₂ compounds decreased (run 2). In contrast, a reactant mixture of the lower O₂/CH₄ ratio gave a better C₂ selectivity but a poorer conversion (run 3). Run 4 shows that a good yield of C₂ was obtained even at a higher partial pressure of the reactants if the O₂/CH₄ ratio was kept at ~0.5. Although C₂H₄ is the main C₂ compound over Li/MgO, C₂H₆ is also produced in considerable amounts. From the practical viewpoint, it is desirable that C₂H₆ be converted into C₂H₄ and the technology for this process is available.

Figure 1 shows effects of reaction temperature on the CH₄ conversion and the selectivities, which were measured over 4 g of 3% Li/MgO. This reaction system is unusual in that the selectivity for C₂ compounds increases as the CH₄ conversion increases up to a temperature of 670 °C. In most heterogeneous oxidation catalysis, the selective oxidation decreases with increasing conversion. The phenomenon in Fig. 1 may be understood in terms of two primary reaction pathways for the $\cdot\text{CH}_3$ radicals; they may either couple to form C₂H₆ (ref. 6) or they may react with the oxide surface to form surface methoxide ions that ultimately result in the production of CO and CO₂ (ref. 7). The former channel is second order with respect to [$\cdot\text{CH}_3$] and is thus favoured at high temperatures where more $\cdot\text{CH}_3$ radicals are produced. One cannot, of course, rule out the possibility that part of the CO and CO₂ is formed by a homogeneous reaction between $\cdot\text{CH}_3$ and O₂. The C₂H₆ may be converted to C₂H₄, also through alkyl radicals, but further oxidation to CO and CO₂ seems to be retarded below 670 °C in the presence of CH₄.

The $\cdot\text{CH}_3$ radicals are believed to be formed at (Li⁺O⁻) centres on MgO (ref. 5) which, have been observed in Li-doped MgO single crystals⁸ and more recently in the sample used in these experiments⁵. Gas phase and surface studies have shown that O⁻ ions efficiently abstract H $\cdot$ from CH₄ (refs 7, 9). The Li/MgO catalyst does not contain ions having variable oxidation states or transition metals; thus, it is a new type of catalyst for the activation of methane.

This work was supported by the NSF. T.I. is on leave from the Department of Chemistry, Tokyo Metropolitan University, Tokyo 158, Japan.

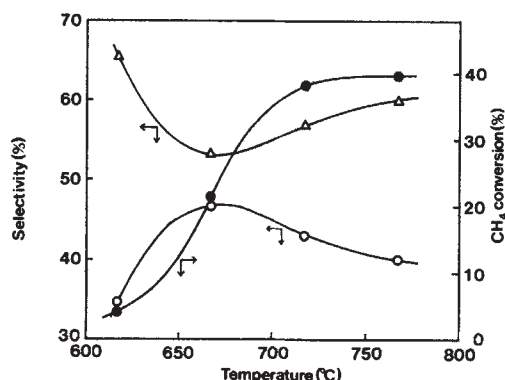


Fig. 1 Effect of the reaction temperature on the partial oxidation of methane. A reactant mixture containing 291 torr of CH₄ and 155 torr of O₂ was fed over 4 g of 3 wt% Li/MgO. Selectivities: ○, C₂ compounds; △, CO plus CO₂; ●, per cent conversion.

Received 27 December 1984; accepted 26 February 1985.

- Keller, G. E. & Bhasin, M. M. *J. Catal.* **73**, 9-19 (1982).
- Atroschenko, V. I., Shchedrinskaya, Z. M. & Gavrya, N. A. *Zh. prikl. Khim., Mosk.* **38**, 643-649 (1965).
- Hinsen, W. & Baerns, M. *Chemikerzeitung* **107**, 223-226 (1983).
- Liu, H.-F., Liu, R.-S., Liew, K. Y., Johnson, R. E. & Lunsford, J. H. *J. Am. chem. Soc.* **106**, 4117-4121 (1984).
- Driscoll, D. J., Martir, W., Wang, J.-X. & Lunsford, J. H. *J. Am. chem. Soc.* **107**, 58-63 (1985).
- Pollard, R. T. *Compreh. chem. Kinet.* **17**, 249-367 (1977).
- Aika, K. & Lunsford, J. H. *J. phys. Chem.* **81**, 1393-1398 (1977).
- Chen, Y., Tohver, H. T., Narayan, J. & Abraham, M. M. *Phys. Rev.* **B16**, 5535-5542 (1977).
- Bohme, D. K. & Fehsenfeld, F. C. *Can. J. Chem.* **47**, 2717-2719 (1969).

Highly conducting acetylene-CO copolymers and limitations of the soliton model

James C. W. Chien, G. N. Babu & J. A. Hirsch

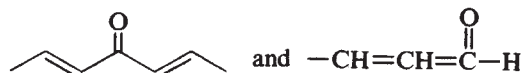
Departments of Chemistry and of Polymer Science and Engineering, Materials Research Laboratories, University of Massachusetts, Amherst, Massachusetts 01003, USA

There is much scientific and technological interest in conductive polymers, of which polyacetylene is the simplest and best example, and in the mechanism of carrier transport. *Trans*-[CH]_x with C_{2h} symmetry is said¹⁻⁴ to have an electronically degenerate ground state with very slight alternation of backbone bond lengths, that is, a high degree of conjugation. Radicals or radical ions in the chain are referred to as soliton^{5,6} and polaron⁷ excitations with moving domains, but these rigorous concepts are quite restrictive and exclusive. For instance, moving-domain excitations cannot exist in *cis*-[CH]_x because much energy is required for the interconversion of *cis*-transoid and *trans*-cisoid structures. However, many physical and spectroscopic properties of undoped and doped *trans*-[CH]_x have been explained by these models⁸. An intersoliton electron-hopping model had been proposed for carrier transport in lightly doped *trans*-[CH]_x⁹. Our purpose here is to test the general validity of the soliton/polaron models. Our present results cast strong doubts on the relevance of these theoretical models to carrier transport in conductive polymers, thereby allowing a search for possible new conductive polymers that is free of the constraints imposed by these models.

Copolymerizations of acetylene and CO were initiated by the Ti(OBu)₄/AlEt₃ catalyst at -78 and 25 °C for 18-24 h at monomer feed ratios (C₂H₂/CO) of 4/1, 2/1 and 1/1 to give copolymers of composition 95/5, 90/10 and 87/13, respectively. The number-average relative molecular masses are ~3,900 ± 300 and 970 for copolymers obtained at -78 and 25 °C, respectively, as determined by quenching with tritiated methanol¹⁰⁻¹².

We have the following evidence that copolymers were

obtained: (1) Elemental analysis (Table 1). (2) Radioassay of ¹⁴CO-containing copolymers. (3) Magic angle spinning ¹³C-nuclear magnetic resonance (NMR) showed a carbonyl resonance at 178 p.p.m.¹³ of the correct intensity. (4) Two carbonyl bands at 1,670 and 1,720 cm⁻¹ attributable¹⁴ to the C=O vibrations in



respectively; the latter was probably formed on quenching of Ti-C(=O)-P with methanol. (5) Copolymers formed directly on electron microscope grids¹⁵ are devoid of the characteristic fibrillar morphology of polyacetylene¹⁶ and are completely amorphous according to electron diffraction^{17,18}. (6) The *cis*-content in the pristine copolymer decreases markedly with the increase of CO monomer; probably, the initially formed *cis*-transoid structures¹⁹ are more easily isomerized to the *trans* structure because of the presence of the single bonds in the -C(=O)- units and the absence of crystallinity. (7) Thermogravimetric analysis of copolymers showed a progressively larger weight loss at lower temperatures with increasing amount of CO, with the weight loss approximately equal to the CO in the copolymer. (8) Pyrolysis GC-MS of the copolymers produced almost the same products as that of polyacetylene²⁰ but with CO and some CO₂ and H₂O. (9) Treatment of the copolymer with protonic acid greatly reduces the intensities of the 178 p.p.m. ¹³C-NMR resonance and the infrared carbonyl bands as expected for facile protonation of carbonyl groups activated by adjacent π systems. Lastly, (10) the copolymer films have a dull coppery colour on the side facing the reactor wall, but the other side has various shiny green colours that depend on the monomer feed ratio and which are different from those of polyacetylene—indicative of the wider band gap for the copolymer than for polyacetylene.

The key magnetic and transport properties of undoped poly(acetylene-co-CO) are summarized in Table 2. For simplicity, we will refer to the samples poly(95acetylene-co-5CO), poly(90acetylene-co-10CO) and poly(87acetylene-co-13CO) as

Table 1 Comonomer feed and copolymer composition

Sample	Mole ratio C ₂ H ₂ /CO in feed	Elemental analysis*			C ₂ H ₂ (mol %)	
		C	H	C ₂ H ₂	By radioassay†	By ¹³ C-NMR‡
A	4	89.43	7.85	95	96	95
B	2	86.34	7.82	88	91	90
C	1	84.90	8.34	85	89	ND

* Microanalysis Laboratory, Massachusetts University; Ti and Al analyses were both 0.2-0.3% as in normal polyacetylene.

† Radioassay (NEN) of copolymers obtained with isotopic carbon monoxide.

‡ ND, not determined.

Table 2 Properties of undoped acetylene-CO copolymers

Sample	<i>cis</i> Content (%)	ΔH_{pp} (G), 25 °C	Spin per repeat unit (+10 ⁵)	σ_{RT} ((Ω cm) ⁻¹)	S_{RT} (μ V K ⁻¹)
' <i>cis</i> '-A*	83	5.5	29	3×10^{-10}	
<i>trans</i> -A†	0	1.1	2.2	4×10^{-5}	
' <i>cis</i> '-B‡	58	5.0	29	1.4×10^{-7}	
<i>trans</i> -B†	0	1.1	8.3	1.7×10^{-6}	1,760
' <i>cis</i> '-C§	37	3.0	4.0	3.5×10^{-5}	1,600
<i>trans</i> -C†	0	1.2	5.0	1.3×10^{-5}	1,640
<i>cis</i> -[CH] _x	~85¶	8-10¶	670-2,000¶		
<i>trans</i> -[CH] _x	<5¶	0.5-1.0¶	6,700-20,000¶		1,450

* Poly(95acetylene-co-5CO); † thermally isomerized from the '*cis*' polymer at 180 °C for 10 min; ‡ poly(90acetylene-co-10CO); § poly(87acetylene-co-13CO).

|| By infrared.

¶ Range of values in literature.

Table 3 Conductivity of doped acetylene-carbon monoxide copolymers

Sample*	$y^{\dagger}(\text{I}_3^-)$	$\sigma_{\text{RT}}((\Omega \text{ cm})^{-1})$	$y(\text{AsF}_5)$	$\sigma_{\text{RT}}((\Omega \text{ cm})^{-1})$	$S_{\text{RT}}(\mu\text{V K}^{-1})$
<i>cis</i> -[CH] _x	0.08	~100	0.12	~400	~+20
<i>p</i> -A‡	0.06	390	0.09	230	+25
<i>p</i> -B	0.05	98	0.03	95	+27
<i>p</i> -C	0.04	129	0.04	161	+18

* Sample composition as in Table 1. † y , Mol. dopant per repeat unit. ‡ *p*, Pristine AC-copolymers.

A, B and C, respectively. The pristine AC-copolymers have room temperature (RT) peak-to-peak linewidths ΔH_{pp} for their unpaired-spin EPR that are comparable to, but smaller than, for *cis*-[CH]_x; the differential is larger for copolymers with more CO content. Their linewidths increased to 6 G at 134 K, consistent with low diffusivity of the unpaired spins. Thermal isomerization of the AC-copolymers greatly reduces the number of unpaired spins to only 10^{-4} – 10^{-3} of the number in *trans*-[CH]_x, but the isomers have comparable RT EPR linewidth and motional narrowing. In sample C, there is on the average one $-\text{C}(=\text{O})-$ unit separating segments containing about six $-\text{CH}=\text{CH}-$ units, the length of which corresponds to the estimated domain widths for soliton²¹ and polaron⁷. The EPR of *trans*-AC-copolymers shows about the same motional narrowing as in *trans*-[CH]_x. This is either because local processes cause line narrowing or because the $-\text{C}(=\text{O})-$ unit does not present a significant barrier to moving domain. Furthermore, the spin-lattice and spin-spin relaxation times are within a factor of two for the two sets of polymers—pristine AC-copolymers and *cis*-[CH]_x and isomerized AC-copolymers and *trans*-[CH]_x. Finally, the undoped AC-copolymers all have conductivity σ_{RT} and thermopower coefficient S_{RT} comparable to the corresponding undoped [CH]_x.

The values of σ_{RT} for iodine-doped pristine AC-copolymers are either comparable to, or greater than, those for iodine-doped *cis*-[CH]_x (Table 3, column 3). The σ_{RT} values of AsF₅-doped AC-copolymers are one-half to one-fourth those of similarly doped *cis*-[CH]_x. The S_{RT} values, indicative of the intrinsic conductivity of the material, are almost the same for AsF₅-doped AC-copolymers and *cis*-[CH]_x (Table 3, column 6).

In conclusion, the AC-copolymers do not have the molecular and supermolecular structures required for soliton and/or polaron excitations; yet those magnetic and transport properties heretofore said to be uniquely belonging to *trans*-[CH]_x and attributed to soliton and/or polaron are manifest in the copolymers. The fact that an extremely low unpaired spin concentration in isomerized AC-copolymers occurs without any loss in conductivity, unlike in *trans*-[CH]_x, invalidates any transport mechanism invoking the participation of neutral soliton. Of course, the situation is very complicated for the heavily doped polymers and much more research has to be done before elucidating further the conduction mechanism. The significance of this work is that the results obviate the rigid and restrictive present theoretical framework for conducting polymers and open up possibilities for creative designing and synthesis of new polymers. Furthermore, it eliminates the need to force the soliton/polaron concept on poly(pyrrole), poly(*p*-phenylene vinylene), poly(*p*-phenylene), poly(phenylene sulphide), poly(*p*-aniline) and others that do not have electronic ground-state degeneracy. A conductive polymer needs to have only short conjugated segments to facilitate the redox doping and stabilize the resulting charge carriers. A similar conclusion was reached by MacDiarmid *et al.*²², who had homogeneously introduced sp^3 -defect sites of the type $-\text{CHD}-$ into polyacetylene and found that the $[-\text{CH}=\text{CH}-]_n\text{CHD}$ material has high conductivity with $\sigma \propto n^{3.2}$ dependence on the conjugation length for $n \geq 3$.

This work was supported by grants from the NSF and the Office of Naval Research. Magic-angle spinning ¹³C-NMR data were obtained in collaboration with Drs P. Bernier and M. Audenaert of the CNRS at Montpellier, France.

Received 24 September 1984; accepted 26 February 1985.

- Shirakawa, H. & Ikeda, S. *Polym. J.* **2**, 231–244 (1971).
- Peierls, R. E. *Quantum Theory of Solids*, 108–110 (Clarendon, Oxford, 1955).
- Chien, J. C. W., Karasz, F. E. & Shimamura, K. *Makromol. Chem. Rapid Commun.* **3**, 655–659 (1982).
- Fincher, C. R. Jr, Chen, C. E., Heeger, A. J., MacDiarmid, A. G. & Hastings, J. B. *Phys. Rev. Lett.* **48**, 100–104 (1982).
- Su, W. P., Schrieffer, J. R. & Heeger, A. J. *Phys. Rev. Lett.* **42**, 1698–1701 (1978).
- Rice, M. J. *Phys. Lett.* **71A**, 152–154 (1979).
- Campbell, D. K. & Bishop, A. R. *Phys. Rev. B* **24**, 4859–4862 (1981).
- Chien, J. C. W. *Polyacetylene—Chemistry, Physics and Material Science* (Academic, New York, 1984).
- Kivelson, S. *Phys. Rev. B* **25**, 3798–3821 (1982).
- Chien, J. C. W., Capistran, J. D., Karasz, F. E., Dickinson, L. C. & Schen, M. A. *J. polym. Sci. polym. Lett. Ed* **21**, 93 (1983).
- Chien, J. C. W., Karasz, F. E., Schen, M. A. & Hirsch, J. A. *Macromolecules* **16**, 1694–1697 (1983).
- Schen, M. A., Karasz, F. E. & Chien, J. C. W. *J. polym. Sci. polym. Chem. Ed* **21**, 2787–2812 (1983).
- Chien, J. C. W. & Babu, G. N. *Macromolecules* (in the press).
- Bellamy, L. J. *The Infra-red Spectrum of Complex Molecules*, 1–132 (Wiley, New York, 1958).
- Karasz, F. E. *et al. Nature* **282**, 286–288 (1979).
- Chien, J. C. W. *et al. Nature* **299**, 608–611 (1982).
- Shimamura, K., Karasz, F. E., Hirsch, J. A. & Chien, J. C. W. *Makromol. Chem. Rapid Commun.* **2**, 473–480 (1981).
- Chien, J. C. W., Karasz, F. E. & Shimamura, K. *Macromolecules* **15**, 1012–1017 (1982).
- Chien, J. C. W., Karasz, F. E., Wnek, G. E., MacDiarmid, A. G. & Heeger, A. J. *J. polym. Sci. polym. Lett. Ed* **18**, 45–52 (1979).
- Chien, J. C. W., Uden, P. C. & Fan, J. L. *J. polym. Sci. polym. Chem. Ed* **20**, 2159–2167 (1982).
- Weinberger, B. R., Ehrenfreund, E., Pron, A., Heeger, A. J. & MacDiarmid, A. G. *J. chem. Phys.* **72**, 4749 (1980).
- Yaniger, S. I., Kletter, M. J. & MacDiarmid, A. G. *188th Am. Chem. Soc. Abstr. Polym.* **154**, (Philadelphia, 1984).

Crustal structure south of the Iapetus suture beneath northern England

M. H. P. Bott, R. E. Long, A. S. P. Green*,
A. H. J. Lewis, M. C. Sinha* & D. L. Stevenson

Department of Geological Sciences, University of Durham,
Durham DH1 3LE, UK

The crustal structure just south of the Iapetus suture¹ along a line extending from the mid-North Sea high across northern England to the Solway Basin has been investigated as part of the Caledonian Suture Seismic Experiment of 1982. This unique sea-to-land wide-angle explosion seismology project has produced data of unprecedented quantity and quality. The main innovation was the use of both closely spaced shots (mainly at sea) and closely spaced stations (mainly on land), yielding a data set which can eventually be arranged in a variety of new ways for phase correlation and modelling, such as constant offset and common mid-point. The line was extended by Irish and German groups across Ireland to the Shannon Estuary, having a total length of over 600 km, but here we present only preliminary results for northern England and adjacent marine regions. The results for this region differ radically from those of the earlier LISPB experiment² in that we have recognized a well-defined 6.15 km s⁻¹ basement interpreted as pre-Caledonian metamorphic basement beneath most of the line at relatively shallow depth, and there is a well-defined mid-crustal refractor starting at approximately 16 km depth beneath the North Sea and northern England which is not recognized beneath the Irish Sea.

* Present addresses: Geothermal Energy Project, Rosemanowes Quarry, Hems, Penryn, Cornwall, UK (A.S.P.G.); Department of Earth Sciences, Bullard Laboratories, Madingley Rise, Madingley Road, Cambridge, UK (M.C.S.)

Fig. 1 Location of Caledonian Suture Seismic Project (CSSP), with details for Irish Sea to North Sea, showing also the supposed Iapetus suture¹ (IS) and the LISPB line².

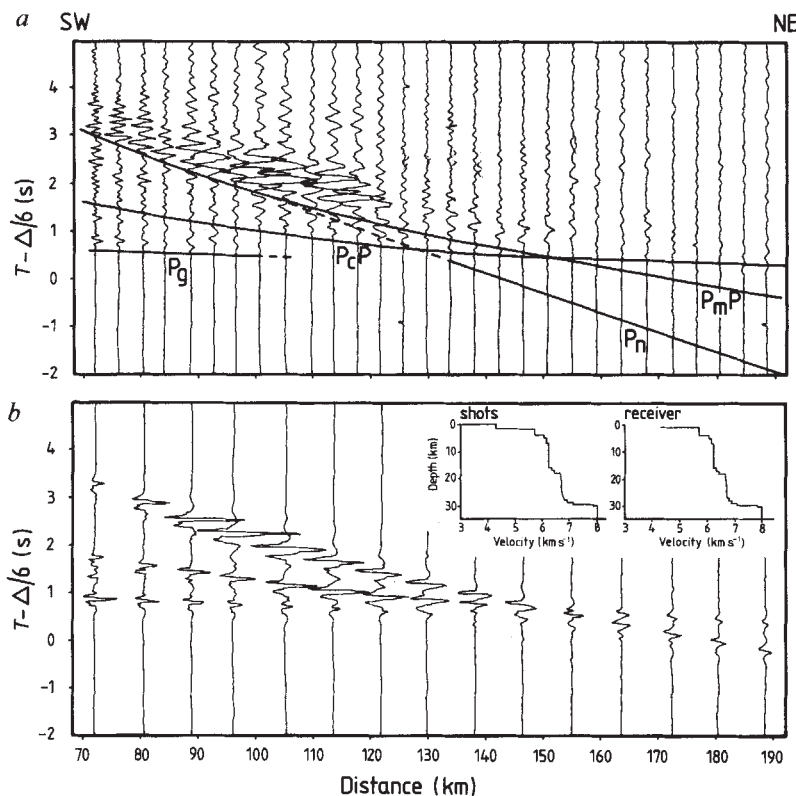
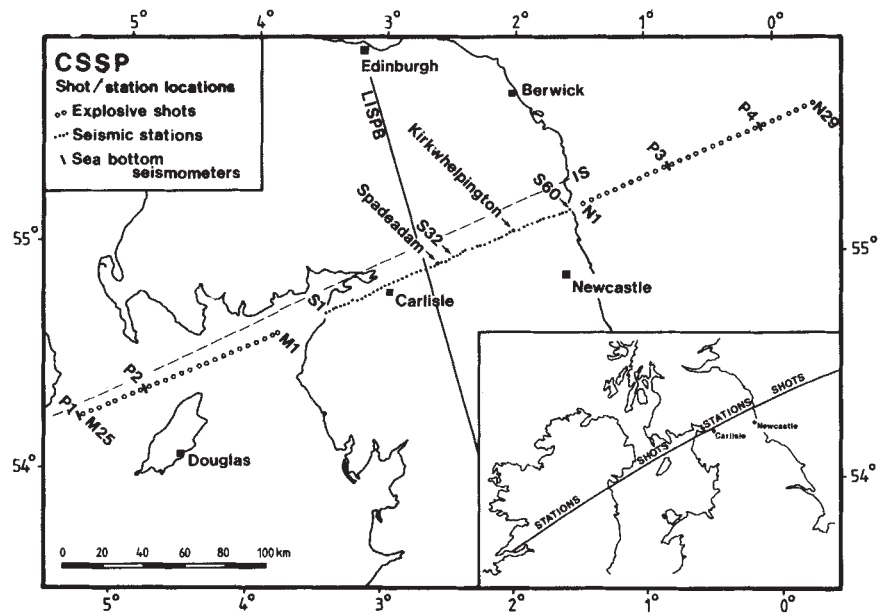


Fig. 2 *a*, Seismic recordings of North Sea shots N1-N29 at station S32 (Fig. 1) reduced to a velocity of 6.0 km s^{-1} , showing main phases computed for the synthetic model. *b*, Synthetic seismogram interpretation of *a* with inset velocity-depth distributions.

The following four main phases have been clearly recognized from the North Sea shots recorded in northern England, as illustrated by the recordings of the shots at station S32 (Figs 1, 2): (1) The first arrival beyond about 30 km is the P_g phase which has a velocity of $6.15 \pm 0.02 \text{ m s}^{-1}$ at 2–4 km depth as determined by time-term analysis. [The symbol P refers to compressional waves. P_g travels in the upper crust. P_cP and P_mP are wide-angle reflections or diving waves from respectively the mid-crustal discontinuity/gradient and the Moho. P_n is the Moho head wave refracted at the crust/mantle boundary.] P_g consistently dies out at all stations for shots beyond about N11–N13. Such decay of the P_g arrival has previously been cited as evidence for low-velocity zones in the upper crust. In the absence of laterally displaced sections containing similar offset distances, it has not been possible in past experiments to distinguish satisfactorily between the effects of a low-velocity zone and effects of lateral velocity variation in the upper crust. In

this case, however, P_g is seen to decay in the vicinity of a specific shot position irrespective of station, so the cause must be interpreted as a lateral inhomogeneity near that shot position rather than as an underlying low-velocity zone. (2) A well-defined P_cP phase of apparent velocity $\sim 6.5 \text{ km s}^{-1}$ results from a steep increase in velocity with depth of $\sim 6.22\text{--}6.5 \text{ km s}^{-1}$ at 16–18 km depth, also detected in some earlier work^{3,4}. (3) The P_mP phase representing the wide-angle reflection from the Moho is a clear large-amplitude arrival. (4) A small amplitude P_n phase of apparent velocity 7.9 km s^{-1} occurs as the first arrival beyond the crossover distance of 130 km. The P_mP and P_n phases are consistent with a sharp Moho at 30 km depth.

A model of the crustal structure assuming horizontal layering beneath the variable shallow structure, together with a synthetic seismogram computed using the reflectivity method⁵, is shown in Fig. 2.

The Irish Sea shots recorded in northern England (Fig. 3)

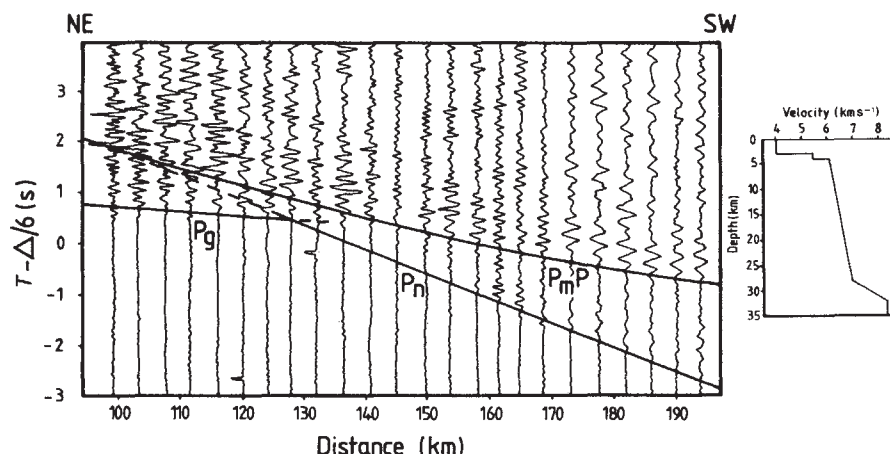


Fig. 3 Seismic recordings of Irish Sea shots M2-M25 at station S32 (Fig. 1) reduced to 6.0 km s^{-1} , showing the main phases computed by ray tracing for inset velocity-depth model.

yielded three main phases: (1) A P_g phase with a velocity of $6.15 \pm 0.02 \text{ km s}^{-1}$ determined by time-term analysis occurs as the first arrival between about 30 km and the crossover distance of 128 km, and comes from a depth of 2–4 km. Arrivals from shots over the Permo-Triassic basin beneath the Solway Firth are systematically delayed by the low-velocity sediments. (2) A conspicuous second arrival of $6.8\text{--}7.0 \text{ km s}^{-1}$ apparent velocity is interpreted as the diving wave from a velocity gradient above the Moho at 28–32 km depth. (3) A P_n phase with an apparent velocity of 8.3 km s^{-1} or higher occurs as the first arrival beyond 128 km. The high apparent velocity is attributed to laterally varying time delays caused by the shallow sediment structure of the Solway basin.

A model of the velocity-depth distribution consistent with these three phases is shown in Fig. 3. A steep increase in velocity marking the top of a lower crustal layer, such as is indicated by the North Sea shots, is not apparent in the Irish Sea. Nevertheless, the possible existence of a higher-velocity lower crust, possibly underlying a velocity inversion at shallower levels, cannot be ruled out at this stage.

The shallow structure along the line of the experiment, which is probably just south of the Iapetus suture, is shown down to the 6.15 km s^{-1} basement in Fig. 4a. The structure down to the base of the Permo-Triassic was not resolved by the experiment and has been obtained from published sources⁶ and from seismic reflection data supplied by Shell. The depth to the 5.5 km s^{-1} Lower Palaeozoic layer has been determined from first arrivals out to ~30 km range. The 6.15 km s^{-1} basement is well defined and its depth, as shown, has been inferred from time-term analysis.

The recognition of this 6.15 km s^{-1} upper crustal refractor at the relatively shallow depth of 4 km or less is a result of considerable significance. This layer extends down to about 16 km depth beneath the North Sea and Northumberland Trough, but its base is undetermined beneath the Irish Sea. The evidence indicates a slight increase in velocity with depth without obvious indication of a low-velocity layer within it. According to Hall *et al.*^{7,8}, velocities $>6.0 \text{ km s}^{-1}$ at shallow depth beneath the Southern Uplands (north of the suture) represent crystalline metamorphic basement beneath the Lower Palaeozoic sequence rather than a metamorphic transition to greenschist facies within the Lower Palaeozoic⁹. Recordings along the Northumberland Trough of a land shot at Spadeadam (Fig. 1) indicate that the boundary between the $5.5\text{--}5.7 \text{ km s}^{-1}$ Lower Palaeozoic rocks underlying the Carboniferous and the 6.15 km s^{-1} layer is sharp rather than gradational, which appears to rule out a metamorphic transition within the Lower Palaeozoic. The 6.15 km s^{-1} velocity is too low for oceanic crust. Both the velocity and the aeromagnetic anomalies¹⁰ along the line are probably too uniform for island-arc-type crust. The simplest interpretation is that the Lower Palaeozoic rocks are relatively thin just south of the Iapetus suture, where they are underlain by crystalline basement which is probably formed of Precambrian metamorphic rocks. The contact could be an unconformity or a tectonic feature. This interpretation differs radically from the LISP profile² crossing our line at Spadeadam, which indicated a $5.8\text{--}6.0 \text{ km s}^{-1}$ Lower Palaeozoic basement extending down to at least 15 km depth.

A second important result is that the deep crust appears to differ between the North Sea/Northumberland Trough region

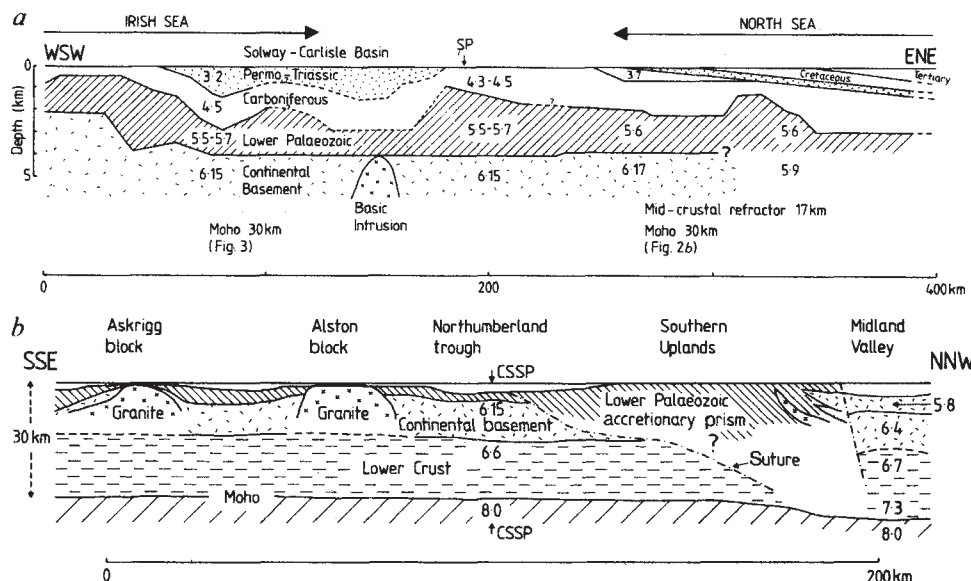


Fig. 4 a, Shallow crustal structure along the line of the experiment (Fig. 1), showing also the approximate locations of the crustal models of Figs 2b and 3. SP, Spadeadam. The basic intrusion beneath the Carlisle basin is inferred from the aeromagnetic map¹⁰. Seismic velocities in km s^{-1} . b, Speculative cross-section perpendicularly crossing the line of the experiment (marked CSSP) from the Midland Valley to the Northern Pennines, showing the possible relationship of the crust south of the Iapetus suture to the Southern Uplands accretionary prism. Seismic velocities are shown in km s^{-1} for this experiment (CSSP) and for the Midland Valley (LISP²).

and the Irish Sea region. P_cP has not been unambiguously detected beneath the Irish Sea, in contrast to the excellent North Sea development, and the Moho is interpreted as being gradational beneath the Irish Sea but relatively sharp beneath the North Sea. Taken at face value, this indicates juxtaposition of different crustal types along strike just south of the suture. This conclusion, however, should be treated with some caution. The low-frequency P_cP and P_mP phases (Fig. 2) are most conspicuously developed for the shots in deeper water in the North Sea, which have significant energy in the 3–5-Hz band. It is possible that the lack of detection of P_cP for the Irish Sea shots, and the weaker P_mP development, may stem from the observed lack of significant energy in this band from the shallower Irish Sea shots. Future work aims to distinguish between these possibilities and to study transition between the two crustal types.

Comparison of our results with the relevant part of the WINCH crustal seismic reflection line¹¹ down the west of Britain is of interest. The WINCH line crosses the western extremity of our Irish Sea shot line, and a short section of WINCH was shot along the western end of our line to allow comparison between the vertical incidence and wide-angle surveys. A conspicuous feature of the WINCH data in the vicinity of our line is the concentration of strong sub-parallel and almost horizontal reflectors which occur at two-way travel times of 5–10 s, contrasting with the much more transparent regions above and below.

The lower limit of these reflectors appears to correspond with the Moho, but from the present preliminary interpretation we have not been able to identify a velocity–depth signature beneath the Irish Sea corresponding to the upper boundary of the zone. It is possible that the distinctive 6.5 km s^{-1} lower crust at the North Sea end of our line corresponds to this zone of reflectors, which is well developed elsewhere on the mainland of England¹², but this needs future testing.

In conclusion, Fig. 4b is a speculative section showing our interpretation of how the crustal structure inferred from this experiment may relate to the Southern Uplands accretionary prism¹³ to the north and the deep structure of the Pennine region⁴ to the south. The velocity–depth distribution beneath the Midland Valley comes from LISPB² and the Iapetus suture itself has been inserted in the form suggested by the WINCH line¹¹ further west. The relationship of the Lower Palaeozoic accretionary prism to the overthrust southern crust explains the observation of a 6.5 km s^{-1} lower crust beneath Eskdalemuir in the Southern Uplands³ and why a higher velocity lower crust was not detected by LISPB².

We thank NERC, the Royal Society, the University of Durham, Mobil and Esso for research grants, Shell for a Studentship (A.H.J.L.) and NERC for shiptime and loan of seismological field equipment. We also thank those who gave field assistance during the project.

Received 6 December 1984; accepted 12 February 1985.

- Phillips, W. E. A., Stillman, C. J. & Murphy, T. *Jl geol. Soc. Lond.* **132**, 579–609 (1976).
- Bamford, D., Nann, K., Prodehl, C. & Jacob, B. *Geophys. J. R. astr. Soc.* **54**, 43–60 (1978).
- Jacob, A. W. B. *Geophys. J. R. astr. Soc.* **18**, 189–197 (1969).
- Bott, M. H. P., Swinburn, P. M. & Long, R. E. *Proc. Yorks. geol. Soc.* **44**, 479–495 (1984).
- Fuchs, K. & Müller, G. *Geophys. J. R. astr. Soc.* **23**, 417–433 (1971).

- Day, G. A. et al. in *Petroleum Geology of the Continental Shelf of North-West Europe* (eds Illing, L. V. & Hobson, G. D.) 76–84 (Institute of Petroleum, London, 1981).
- Hall, J. et al. *Nature* **305**, 418–420 (1983).
- Hall, J. et al. *Nature* **309**, 89–90 (1984).
- Oliver, G. J. H. & McKerrow, W. S. *Nature* **309**, 89 (1984).
- 1:250 000 Series Aeromagnetic Anomaly Maps (Institute of Geological Sciences, London).
- Brewer, J. A. et al. *Nature* **305**, 206–210 (1983).
- Whittaker, A. & Chadwick, R. A. *Geol. Mag.* **121**, 621–624 (1984).
- Leggett, J. K., McKerrow, W. S. & Eales, M. H. *Jl geol. Soc. Lond.* **136**, 755–770 (1979).

Hydrothermal manganese plumes in the Mid-Atlantic Ridge rift valley

G. Klinkhammer*, P. Rona†, M. Greaves* & H. Elderfield*

* Department of Earth Sciences, University of Cambridge, Cambridge CB3 0EZ, UK

† NOAA/AOML, 4301 Rickenbacker Causeway, Miami, Florida 33149, USA

The concentration of dissolved manganese in sea water can be used as a chemical tracer of hydrothermal activity on the underlying sea floor. Here we present vertical profiles of seawater manganese concentrations above the rift valley of the Mid-Atlantic Ridge (MAR). The values obtained indicate the presence of plumes that are presently active generated by hydrothermal vents in the rift valley. These plumes are confined to the rift valley and do not spill over into the adjacent deep ocean basins. Differences between them suggest that there are at least five hydrothermal sources along this 1,700-km-long section of mid-ocean ridge.

The 9 MAR sites (Fig. 1) were preselected on the basis of various indicators of either former or present hydrothermal activity including a magnetic signature characteristic of hydrothermal alteration of the basaltic layer of oceanic crust¹, excess ³He in the water column and hydrothermal mineralization on the sea floor^{2–3}. Manganese has been established as a sensitive tracer of hydrothermal emanations^{4–8}.

The analyses were carried out as part of an extensive regional survey by the NOAA ship *Researcher* of the MAR between the TAG site at 26°N and the Vema Fracture Zone at 11°N in August 1984. Manganese was measured on board ship by flameless atomic absorption after preconcentration with oxine⁹ to give a detection limit of $0.05 \text{ nmol kg}^{-1}$, four times lower than manganese concentrations normally found in deep-ocean waters¹⁰. A portable Zeeman tungsten furnace atomic absorption

spectrophotometer (model AAZ, Scintrex Corporation) was used.

Most manganese data for ocean waters have been reported as 'total dissolvable manganese' (TDM); that is, the amount of manganese measured in an unfiltered sample after acidification to pH 2 with hydrochloric acid¹¹. TDM represents essentially all the manganese, both dissolved and particulate, except for possibly the most refractory component⁹. Selected samples were analysed for TDM. However, TDM analyses are unsuitable for survey work because samples must be stored at pH 2 to ensure that leaching of particles is complete. To overcome this problem, samples were analysed without prior acidification. In this way, six samples could be analysed within <2 h after the samples had arrived on deck. The manganese measured before acidification will be referred to as 'total reactive manganese' or TRM.

Samples were collected in 5-l Niskin and 30-l Go-Flo bottles mounted on a rosette or strung on a hydrowire. Fourteen aliquots including eight from plumes in the rift valley were filtered before being analysed for manganese. The average coefficient of variation between the concentrations of these filtered aliquots and the corresponding TRM results was 3.8%. This exercise demonstrates that TRM represents mainly the dissolved fraction as defined by passing through $0.4\text{-}\mu\text{m}$ filters. The difference between TDM and TRM is the manganese released from particles by acidification to pH 2 or acid-soluble manganese (ASM). In deep Atlantic waters not affected by hydrothermal discharge, excess particulate manganese is $<0.04 \text{ nmol kg}^{-1}$ (ref. 12), below the detection limit of the method used here. Therefore, in deep waters not affected by hydrothermal activity or resuspension, TDM and TRM should be equivalent within our analytical precision, which they were at the control station (Fig. 2a).

Typically, a manganese-depth profile of the open ocean is similar to the one at Vents station 1 (Fig. 2a) over abyssal hills some 660 km west of the MAR (26°00.0' N; 51°00.0' W). This profile, a control with which to compare results from the rift valley, consists of a surface water maximum of $1.63 \text{ nmol kg}^{-1}$ and deep-water values of $0.26 \pm 0.08 \text{ nmol kg}^{-1}$; these results are

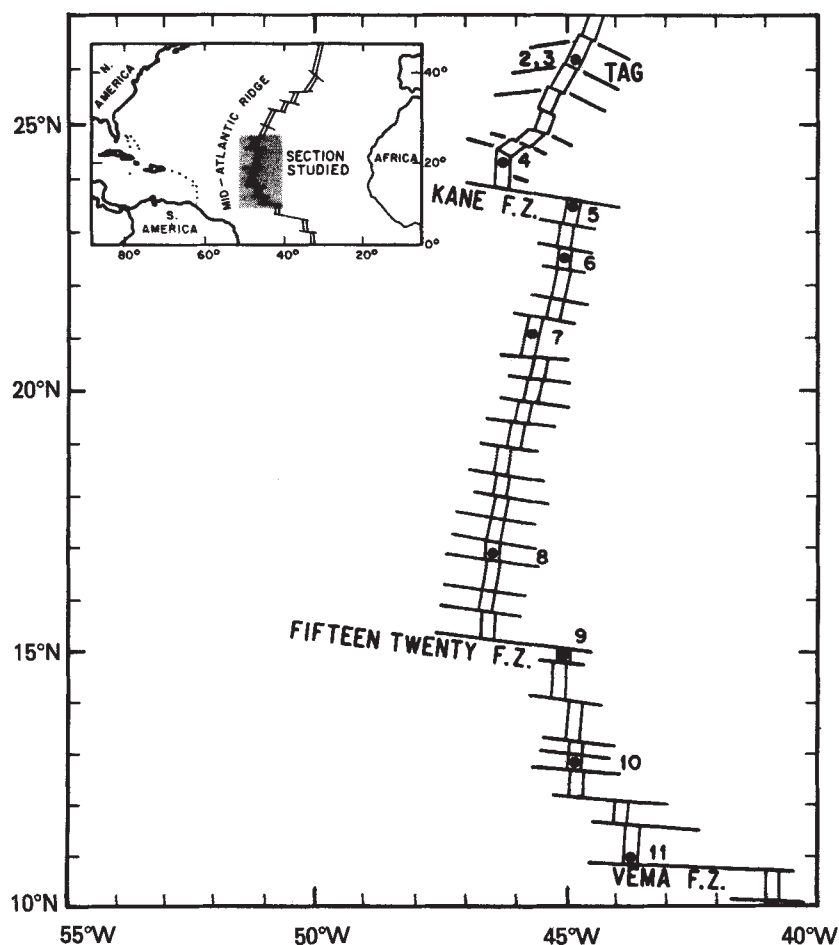


Fig. 1 Study area. Water sampling during the survey consisted of 11 stations: station 1, the control station west of the spreading centre, and stations 2-11 in the median valley of the MAR.

similar to data published for the nearby Nares Abyssal Plain¹³. Input from aeolian and riverine sources result in surface-water maxima at northern latitudes, but oxidation and scavenging maintain much lower concentrations in deep waters. Three processes may produce manganese anomalies below 1,000 m at open ocean sites: (1) diffusion from sediment after diagenetic remobilization; (2) resuspension of sediment by bottom currents; or (3) discharge of hydrothermal effluents. Diffusion from sediments is unlikely to be significant at our sites for two reasons. First, most deep-sea sediments are capped by an oxidizing layer that traps remobilized manganese within the sediment¹⁴. Second, there is very little sediment accumulation in the rift valley. Process (2) may produce an anomaly near the bottom, but such a signal is associated with particulate matter and thus would result in an anomaly in TDM but not TRM. In conclusion, the manganese anomalies discovered during this survey are most certainly produced by process (3).

Manganese is highly enriched in hydrothermal effluents because it is leached from basalt during hydrothermal circulation. The concentration of manganese in a hydrothermal end-member is $\sim 1 \text{ mmol kg}^{-1}$ (refs 15, 16), five million times greater than in deep ocean waters. Figure 2b is a TDM-depth profile over the vent field at Galapagos⁴. The buoyant hot-spring effluents emanating from the Galapagos Spreading Centre create a manganese anomaly that extends from the sea bed some 1,000 m into the water column. In contrast, the profile above the 'black smokers' at 21°N contains a manganese maximum 220 m above the sea floor (Fig. 2c)⁶.

As well as the size and shape of the anomaly, the partitioning of manganese between dissolved and particulate phases provides useful information. Determining partitioning is useful because the manganese in reducing hydrothermal fluids is dissolved, but once discharged into oxygenated sea water, it converts to the particulate form as a result of oxidation and scavenging. Hence, the proportion of manganese present as particles is an indication

of the age of the hydrothermal signal. Here, the difference between TDM and TRM (that is, ASM) will be used to infer distance from a source. However, we note that ASM is a function of time and not distance and, because both the kinetics of ASM formation and the directions of currents at our sites are unknown, such deductions should be considered tentative. Our results are presented in Table 1 and Fig. 3a-i.

Stations 2, 3 (TAG; Fig. 3a): Previous evidence for hydrothermal activity at the TAG site was substantial. The area is characterized by rapidly accumulating manganese crusts¹⁷ and high metal contents in sediments¹⁸; temperature¹⁹ and helium²⁰ anomalies in bottom waters have also been reported.

The manganese results from the TAG site show that there are active hydrothermal vents on the MAR, a conclusion supported by results from other sites occupied during the survey. Whether or not there are active vents at the TAG site is less clear. The manganese anomalies at station 2 are consistent with the presence of hydrothermal emanations from the east wall, as concluded from previous work at TAG²¹. If the anomaly is maintained by local activity, the largest sources probably occur at 3,450 m, somewhat deeper than previously thought. However, the results from TAG do not eliminate the possibility that the manganese anomaly here is advected from elsewhere along the MAR and that there is no present-day venting at this site. Indeed, the observation of large amounts of particulate manganese at station 2, L2, would support such a conclusion.

Station 4 (Fig. 3b): It is irrefutable that the excess manganese at this site is the result of hydrothermal activity. The shape of this profile is reminiscent of those found in the South Pacific⁸ and could result from buoyant emanations from the valley floor, advection from the valley wall(s) near these locations or advection from vents some distance away. The small concentrations encountered and the observation that >20% of the anomalous TDM is associated with suspended particulates suggests that the anomaly here is an advected feature.

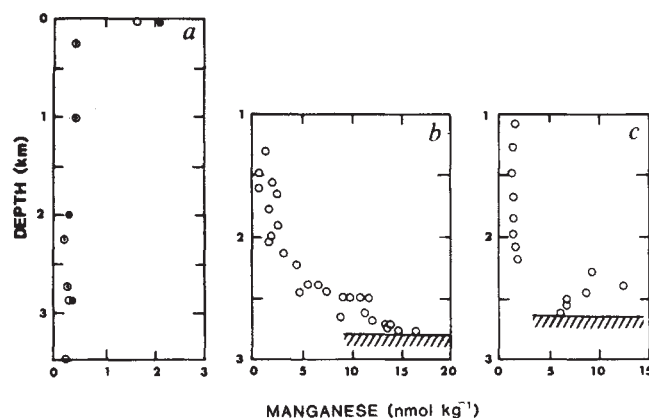


Fig. 2 a, Manganese-depth profile at the control station. Solid circles are total reactive manganese (TRM) and open circles are total dissolvable manganese (TDM). Circles with dots indicate identical concentrations within precision. The bottom depth is ~5,000 m. b, TDM profile over the hot springs on the Galapagos Spreading Centre (redrawn from ref. 4). c, TDM profile over the 'smokers' on the East Pacific Rise at 21° N (redrawn from ref. 6).

The hatched lines in b and c indicate the bottom depth.

Station 5 (Fig. 3c): This station is in the rift valley at its eastern intersection with the Kane Fracture Zone (Fig. 1). Again, evidence for present-day hydrothermal activity was found. Whereas the size, shape and position of the anomaly here are similar to those at station 4, the form of the manganese is distinctly different; that is, there is very little particulate manganese (ASM) at this station, suggesting closer proximity to a source.

Station 6 (Fig. 3d): There is a distinct manganese anomaly at this station and very little ASM, indicating close proximity to active vents. The deepest sample has more particulate manganese, which may be explained by resuspension of sediments. The most obvious difference between the profile at station 6 and those further north is that the maximum concentration occurs much shallower, at 3,000 m as opposed to 3,500 m (although there is a small anomaly at 3,500 m as well).

Station 7 (Fig. 3e): The TRM anomaly at this location is the smallest found in the rift valley. There is 0.2–0.3 nmol kg⁻¹ of particulate manganese in most samples as well, which may be related to the fact that the axis of the rift valley is anomalously offset ~30 km to the west at this site.

Station 8 (Fig. 3f): The TRM profile at station 8, L2, was obtained by positioning the rosette over the east wall at a depth

as near as possible to 3,800 m. The most probable explanation for these results is that the rift valley anomaly (station 8, L1) is produced by advection of hydrothermal manganese from hot springs on the east wall at this site and that this venting occurs at a depth of 3,800 m. This tentative conclusion is based on four observations: (1) the maximum TRM concentration in the valley is 1.90 nmol kg⁻¹, compared with 2.58 nmol kg⁻¹ over the wall; (2) these maxima occur at approximately the same depth in both profiles; (3) there is much more ASM in the valley, indicating that the signal there is older; (4) the shape of the east wall profile is similar to the one obtained over hot springs on the Galapagos Spreading Centre (Fig. 2b).

Station 9 (Fig. 3g): The results from station 8 suggested that the venting there may be asymmetrical about the spreading axis, that is, occurring on the east wall but not the west wall. To examine this problem more closely, two rosette lowerings were made at station 9. These results demonstrate that the distribution of hydrothermal manganese is asymmetrical at this site. Because the manganese plumes contain large quantities of ASM, both locations at this site may be somewhat removed from the nearest vents. The ASM results are similar to the concentrations of 'excess' manganese (0.6–1.0 nmol kg⁻¹) found by Lambert *et al.*¹² in the rift valley at 14°15' N. The shapes and relative sizes of the anomalies suggest that the nearest venting is associated with the east wall and that the anomaly on the west wall is produced by westward advection.

Station 10 (Fig. 3h): The maximum in the valley occurs at about 3,600 m and is larger than the maximum over the east wall. These observations, together with the presence of more suspended particulate manganese at the east wall location, makes westward advection similar to that deduced for station 9 improbable. Indeed, the results are more compatible with a west wall source at this site.

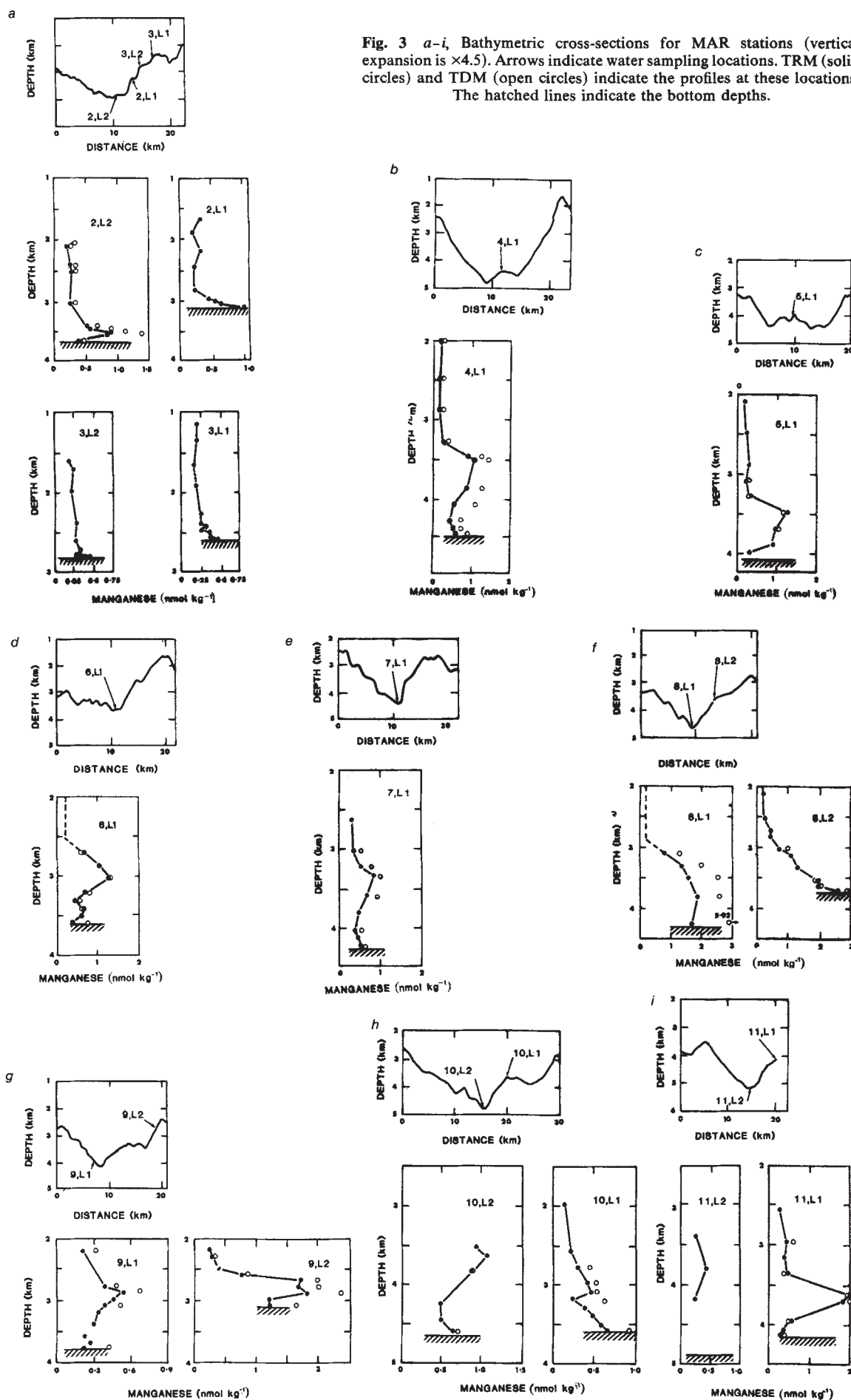
Station 11 (Fig. 3i): This station was positioned in the rift valley at its western intersection with the Vema Fracture Zone. The TDM results demonstrate that the anomaly over the east wall is essentially all dissolved manganese. Given this evidence, station 11 must be considered as a promising site for more detailed investigation.

We conclude from the manganese plumes discovered during this survey that there are active hydrothermal vents along this section of the Mid-Atlantic Ridge. The differences between these plumes make it virtually impossible for them to have originated at the same vent field. First, they occur at different depths (2,900–3,800 m), over different ridge segments (Fig. 1) offset by as much as 150 km. Furthermore, the size and shape of the anomaly change, both across the rift and along the spreading

Table 1 Manganese results for MAR rift valley stations

Station	Lowering	Location	Bottom depth (m)	TRM maximum*				
				Depth (m)	TRM	TDM (nmol kg ⁻¹)	ASM	% + RM
2	1	26°09.7' N; 44°48.1' W	3,130	3,090	1.02	—	—	—
	2	26°09.7' N; 44°46.6' W	3,674	3,468	0.91	1.16	0.25	78
3	1	26°08.0' N; 44°47.5' W	2,617	2,597	0.46	—	—	—
	2	26°08.5' N; 44°46.5' W	2,800	2,786	0.47	—	—	—
4	1	24°22.5' N; 46°14.2' W	4,480	3,504	1.07	1.41	0.34	76
5	1	23°27.8' N; 44°55.5' W	4,075	3,475	1.28	1.15	—	100
6	1	22°28.5' N; 45°03.3' W	3,626	3,026	1.31	1.30	—	100
7	1	21°06.0' N; 45°46.5' W	4,232	3,324	0.83	0.93	0.10	89
8	1	16°47.7' N; 46°26.6' W	4,300	3,800	1.90	2.62	0.72	73
	2	16°48.0' N; 46°23.0' W	3,705	3,685	2.58	2.84	0.26	91
9	1	14°53.8' N; 45°01.6' W	3,798	2,873	0.54	0.67	0.13	81
	2	14°55.0' N; 44°54.0' W	3,079	2,879	1.80	2.35	0.55	77
10	1	12°55.0' N; 44°49.0' W	3,590	3,574	0.69	0.95	0.26	73
	2	12°55.3' N; 44°50.5' W	4,645	3,620	1.18	—	—	—
11	1	11°01.3' N; 44°38.1' W	4,165	3,627	2.02	2.00	—	100
	2	11°01.5' N; 44°42.0' W	4,992	—	0.67	—	—	—

* TRM, Total reactive manganese; TDM, total dissolvable manganese; ASM, acid-soluble manganese.



axis. Finally, the proportion of the excess manganese associated with particles is variable.

By taking a conservative approach and grouping those plumes in which the maxima occur at the same depth, five sources can be identified, that is stations 2, 4 and 5; stations 6 and 7; station 8; station 9; and stations 10 and 11. This analysis predicts a frequency of one source per 340 km of ridge; of course the actual frequency may be much higher.

The position of venting in the rift appears to vary with latitude. Results from stations 8 and 9 are consistent with venting from the east wall, whereas results from station 10 indicate a west-wall source. Furthermore, these results are consistent with discharge from the valley walls, but they do not eliminate the possibility that venting occurs from the valley floor. Finally, hydrothermal manganese along this section appears to be confined

to the median valley and does not spill over into the adjacent ocean basins, although it may leak out through fracture zones.

The anomalies reported here are smaller than those found near venting in the Pacific⁴⁻⁷. However, this difference is an artefact of the regional sampling strategy of this survey and is not a true indication of the relative impact of venting from the MAR on geochemical cycles and the heat budget. Such information can be acquired only by site-specific work.

We thank Captain R. L. Swanson, officers and crew of the NOAA ship *Researcher* for their support and T. A. Nelsen of NOAA and J. Trefry of Florida Institute of Technology for their cooperation. This work was supported by the NOAA Vents programme and the NERC; Cambridge University Earth Sciences Series, contribution 578.

Received 21 November 1984; accepted 21 February 1985.

1. Rona, P. A. *J. Volcanol. geotherm. Res.* **3**, 219-225 (1978).
2. Rona, P. A. *et al. Geol. Soc. Am. Abstr. Progr.* **14** (7), 602 (1982).
3. Rona, P. A. *Earth Sci. Rev.* **20**, 1-104 (1984).
4. Klinkhammer, G., Bender, M. & Weiss, R. F. *Nature* **269**, 319-320 (1977).
5. Klinkhammer, G. P. *Chem. Geol.* **29**, 211-226 (1980).
6. Lupton, J. E. *et al. Earth planet. Sci. Lett.* **50**, 115-127 (1980).
7. Jones, C. J., Johnson, H. P. & Delaney, J. R. *Geophys. Res. Lett.* **8**, 873-876 (1981).
8. Klinkhammer, G., Elderfield, H. & Hudson, A. *Nature* **305**, 185-188 (1983).
9. Klinkhammer, G. *Analyt. Chem.* **52**, 117-120 (1980).
10. Klinkhammer, G. P. & Bender, M. L. *Earth planet. Sci. Lett.* **46**, 361-384 (1980).
11. Bender, M., Klinkhammer, G. & Spencer, P. *Deep-Sea Res.* **24**, 799-812 (1977).
12. Lambert, C. E., Bishop, J. K. B., Biscaye, P. E. & Chesselet, R. *Earth planet. Sci. Lett.* **70**, 237-248 (1984).

13. Burton, J. D., Maher, W. A. & Statham, P. J. in *Trace Metals in Seawater* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 415-426 (Plenum, New York, 1983).
14. Froelich, P. N. *et al. Geochim. cosmochim. Acta* **43**, 1075-1090 (1979).
15. Bischoff, J. L. & Dickson, F. W. *Earth planet. Sci. Lett.* **25**, 385-397 (1975).
16. Edmond, J. M., Von Damm, K. L., McDuff, R. E. & Measures, C. I. *Nature* **297**, 187-191 (1982).
17. Scott, M. R., Scott, R. B., Rona, P. A., Butler, L. W. & Nalwalk, A. J. *Geophys. Res. Lett.* **1**, 355-358 (1974).
18. Shearman, S., Cronan, D. S. & Rona, P. A. *Mar. Geol.* **51**, 269-291 (1983).
19. Rona, P. A. *Geophys. Res. Lett.* **5**, 993-996 (1978).
20. Jenkins, W. J., Rona, P. A. & Edmond, J. M. *Earth planet. Sci. Lett.* **49**, 39-44 (1980).
21. Rona, P. A. *J. geol. Soc., Lond.* **137**, 385-402 (1980).

Fractal dimension of vegetation and the distribution of arthropod body lengths

D. R. Morse*, J. H. Lawton*, M. M. Dodson† & M. H. Williamson*

Departments of *Biology and †Mathematics, University of York, York YO1 5DD, UK

Following Mandelbrot¹, recent studies²⁻⁶ demonstrate that some natural surfaces are fractal. Here we show that transects across vegetation are fractal, and consider one possible consequence of this observation for arthropods (mainly insects) living on plant surfaces. An important feature of a fractal curve or surface is that its length or area, respectively, becomes disproportionately large as the unit of measurement is decreased¹. This suggests that if vegetation has a fractal structure, there is more usable space for smaller animals living on vegetation than for larger animals. Hence, there should be more individuals with a small body length than a large body length. We show that this is the case, and that relative numbers of small and large individual arthropods collected from vegetation are broadly consistent with theoretical predictions originating from the fractal nature of vegetation⁷ and individual rates of resource utilization.

Mandelbrot^{1,4} asks, how long is the coast of Britain? Measuring its length by stepping round with successively shorter divider step-lengths allows finer and finer resolution of the coastline; if the log of divider step-length is plotted against the log of total length measured, a straight line of negative slope is obtained^{1,4,8}. Mandelbrot^{1,4} interprets this as evidence that coastlines are fractal and that the slope of the graph is an estimate of $1-D$, where D is the fractal dimension of the coastline. The relationship between D , step-length r and the perceived length of the line $l(r)$ measured with step r is given by

$$l(r) \propto r^{(1-D)} \quad (1)$$

By definition¹, for a line (that is, a transect across a surface), D must lie in the range $1 \leq D \leq 2$ and for a surface, $2 \leq D \leq 3$.

The fractal dimension of vegetation might also be measured by stepping along a transect over the surface with a pair of dividers set to a series of step-lengths and using equation (1). This was the method used to measure the fractal dimension of

transects across a coral reef⁶. An alternative widely used method (see ref. 1, p. 130), which has some practical advantages, is as follows.

Enclose the object (or a subsection of it) in a square of side S and divide the square into $(S/\lambda)^2$ squares of side λ . Let $N(\lambda)$ be the number of squares the edge of the object enters and plot $\log N(\lambda)$ against $\log 1/\lambda$. Suppose that for small values of λ , the graph is almost linear with slope D . Then D can be interpreted as the fractal dimension, since

$$\log N(\lambda) \approx D \log 1/\lambda + \log K \quad (2)$$

where K is a constant, whence

$$N(\lambda)\lambda^D \approx K \quad (3)$$

The method is explained in greater detail in Fig. 1 legend, together with an example.

Photographs were taken of a variety of plants during early spring. The plants were either leafless, just coming into leaf or evergreen. By careful choice of plant-parts and camera focal length, the resulting photographs were approximately two-dimensional, greatly simplifying their interpretation under the grid (Fig. 1a). Two different arbitrarily sized grids (128 mm or 256 mm along one edge) were partitioned into 2^n squares along each edge. Depending on the grid size, n varied from 2 to 6 or 7, down to a square size of 2 mm. The grid was randomly repositioned several times to give semi-independent estimates of the fractal dimension of the same piece of vegetation. The slope of the resulting line (for example, Fig. 1b) was estimated by regression analysis.

There are several points to note in the interpretation of such graphs. One is that when the number of divisions is small, it is probable that all the squares will be entered, hence the slope of the graph will be $D=2$. This is an artefact of the method; consequently, in the present study all points that fell on the line $y=2x$ (on a log/log plot) were omitted when estimating the slope of the line. At the other extreme of resolution, when a very large number of squares is used, the slope of the graph will in general fall to 1. This could be due to a lack of resolution in the photograph or because irregularities in the outline of the plant no longer occur at that scale. Taking a photograph at a much larger scale (higher magnification) resolves this problem. The graph may be convex^{9,10} between the extremes of $D=2$ and $D=1$, or it could have two straight subregions characterized by different slopes (Fig. 1b). Similar changes in fractal

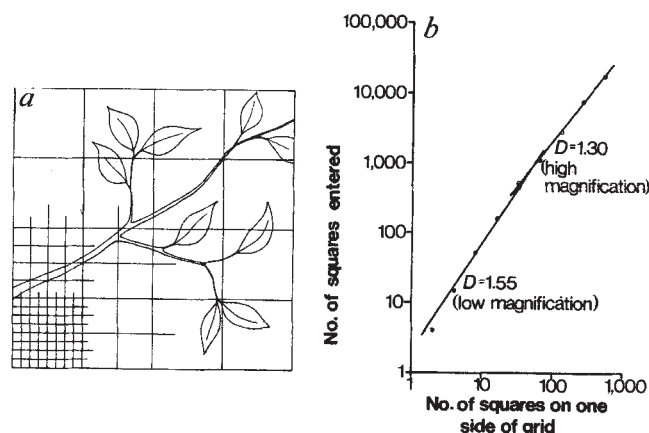


Fig. 1 *a*, Photographs of plants at various magnifications were placed under a grid. The number of squares entered by the outline of the plant were counted, starting with a coarse grid of two large squares on one side, then 2^n squares, with n varying from 2 to 6 or 7, depending on the grid size. For ease of representation, the plant's leaves in this figure are drawn flat; in reality they are orientated at all angles with respect to the grid. Also for clarity, the progressively finer divisions are only illustrated in one corner of the figure. The logarithm of the number of squares entered by the outline of the plant was then plotted against the logarithm of the number of squares along one side of the grid, as in *b*. The slope of the line equals the fractal dimension, D . *b*, Data gathered in this way for Virginia creeper, photographed without leaves in early spring. The twigs were photographed at one scale, then parts of the same twigs were rephotographed at a higher magnification, permitting D to be estimated at two levels of resolution.

dimension at different scales have been found in other natural objects^{1,3,11}.

Table 1 presents data on the fractal dimension of a selection of plants. Estimates of D range from 1.28, for a close-up photograph of Virginia creeper, to 1.79 for cotoneaster. Those plants which one might, *a priori*, consider to have a more complex growth form have a higher fractal dimension. In those plants which were photographed at two scales, estimates of the fractal dimension are lower at the higher magnification, as would be expected from the above argument.

The mean fractal dimension of the samples in Table 1 is 1.44. For ease of calculation, assume that, typically, $D \approx 1.5$. It follows from equation (1) that for an order-of-magnitude decrease in

ruler length, the expected distance between two points on a linear transect (of dimension $D = 1.5$) across a fractal surface increases by a factor of $\sqrt{10} = 3.16$. The difficulties of measuring the fractal dimension of a surface are considerable and no data are yet available. However, squaring the increase in linear distance (that is, adding the fractal dimensions of the orthogonal transects; ref. 1, p. 365) gives a heuristic upper bound estimate for the expected increase in surface area. Therefore, the maximum expected increase in surface area is $3.16^2 = 10.0$ for an order-of-magnitude decrease in ruler length. Bearing in mind the disconnected character of the surface of vegetation, an estimate for the lower bound of the fractal dimension is obtained by adding 1 to the linear fractal dimension (for example, $1.5 + 1$; see ref. 1, p. 365). This holds exactly under some circumstances, and in particular when the surface is flat in a direction transverse to the cross-section (which is clearly not the case for vegetation). This lower bound predicts a 3.16-fold increase in surface area for an order-of-magnitude decrease in ruler length, when $D = 1.5$.

Now consider the implications of the fractal nature of plant surfaces for the animals living on them. As a first approximation, substitute animal body length (L) for step-length (λ) in equation (1). If the way in which animals perceive and use their environment is proportional to their body length¹², then for a homogeneous fractal surface having transects with $D = 1.5$, the area perceived by animals 3 mm long may be up to an order of magnitude greater than the area perceived by animals 30 mm long, for the same reference area. This increase in available space for animals of smaller body length may be combined with a consideration of the way in which metabolic rate scales with body length^{13,14} to make predictions about the distribution of body lengths of animals living on vegetation.

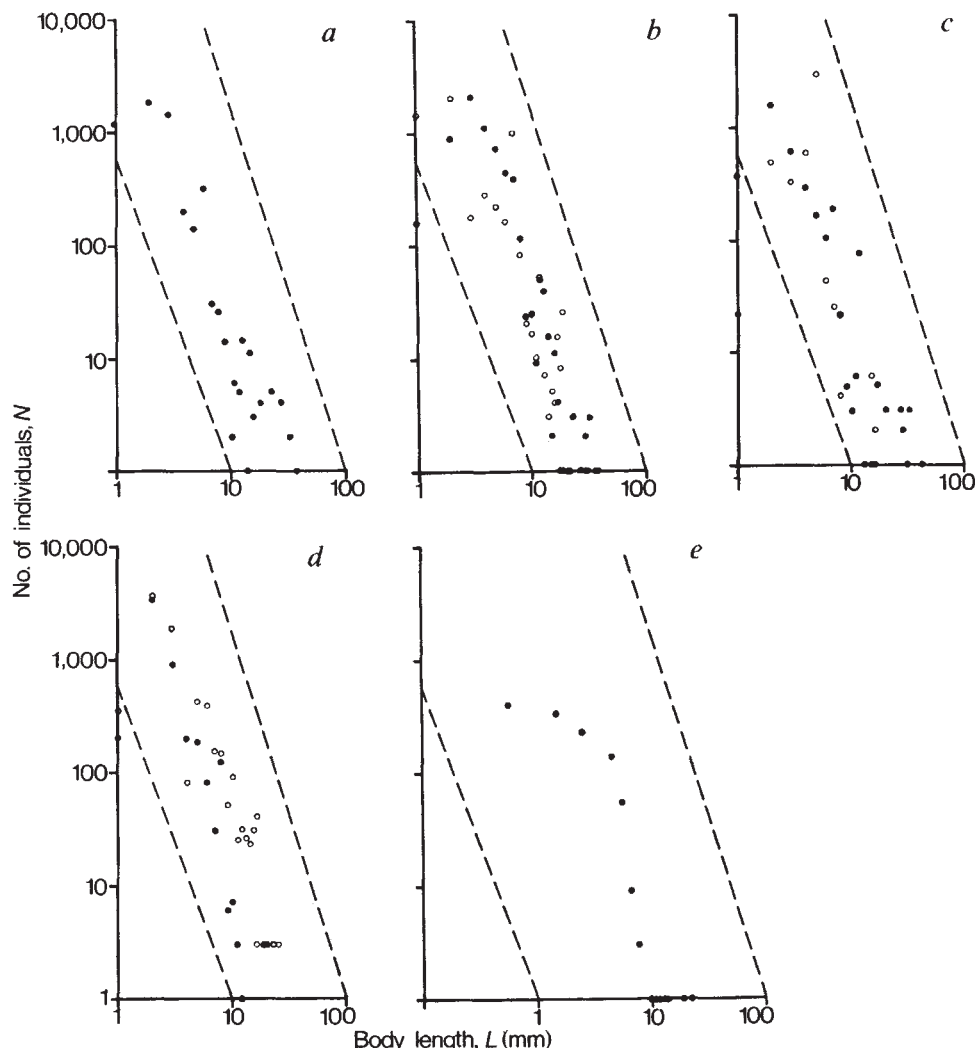
The metabolic rate of individual animals^{13,14} scales approximately as the 0.75 power of body weight, W , that is, as $(L^3)^{0.75}$. Next, suppose that population densities are approximately proportional to the reciprocal of individual rates of resource utilization (that is, to metabolic rate⁻¹; see, for example, ref. 13). Then, if use of resources per individual is proportional to $W^{0.75}$, it follows that population density, N , scales as $(L^3)^{-0.75}$. Hence, all other things being equal (especially the rate of appearance of new resources), a 10-fold decrease in body length results in a $(10^3)^{0.75} = 178$ -fold increase in the density of individuals. This increase in density may be combined with the expected increase in the available surface area, outlined above, for a given decrease in body length, to predict relative numbers of individual animals of different body lengths living on the surface of vegetation. For an order-of-magnitude decrease in body length, such calcula-

Table 1 Estimate of the fractal dimension, D , of woody plants for arbitrarily positioned transects across their surfaces

Species	Magnification	No. of estimates	Mean D	s.d.
Barberry, <i>Berberis vulgaris</i> L. (evergreen)	High	3	1.46	0.018
	Low	3	1.43	0.042
Virginia creeper, <i>Parthenocissus tricuspidata</i> (Sieb. and Zucc.) Planch. (twigs and buds)	High	6	1.28	0.078
	Low	3	1.55	0.009
Weeping elm, <i>Ulmus glabra</i> forma <i>pendula</i> (Loud.) Rehd. (twigs and buds)	Low	6	1.41	0.111
Cotoneaster, <i>Cotoneaster horizontalis</i> Decaisne (twigs and leaves)	High	3	1.35	0.019
	Low	6	1.79	0.093
Ivy, <i>Hedera helix</i> L. (evergreen)	Low	18	1.39	0.050
Yew, <i>Taxus baccata</i> L. (evergreen)	High	3	1.47	0.042
	Low	3	1.68	0.099
Silver birch, <i>Betula pendula</i> Roth (twigs and leaves)	Medium	3	1.40	0.040
Downy birch, <i>Betula pubescens</i> Ehrh. (twigs and leaves)	Medium	3	1.40	0.035
Ash, <i>Fraxinus excelsior</i> L. (twigs and leaves)	Medium	3	1.42	0.083
Sycamore, <i>Acer pseudoplatanus</i> L. (twigs and leaves)	Medium	3	1.31	0.023

Photographs of branches and twigs of a selection of woody plants, with or without leaves, were taken from the University of York campus or Skipwith Common, North Yorkshire. High-magnification estimates were derived from close-up photographs of twigs ~ 10 cm long; low-magnification estimates were from photographs of twigs > 50 cm long. Single medium estimates were made from specimens > 25 cm long. The state of each plant when photographed and the plant part photographed is indicated in parentheses after the species name. (Note that sample transects inevitably tended to follow twigs and other 'principal axes' in the natural structure of the vegetation. This could introduce some bias into the estimates of their fractal dimension.)

Fig. 2 Data on the number of individual arthropods (mainly insects) of different body lengths collected from vegetation. *a*, Understorey foliage in primary forest, Finca Tabogo, Costa Rica¹⁵; *b*, Osa secondary vegetation (●) and Kansas secondary vegetation (○); *c*, Tabogo primary riparian vegetation (●) and Icos vegetation (○)¹⁶; *d*, understorey foliage in cacao plantations in Dominica (●) and at Finca La Lola, Costa Rica (○)¹⁷; *e*, Birch (*Betula pubescens*) trees at Skipwith Common, North Yorkshire. In *a-d*, arthropods were collected by sweep net, in *e* by pyrethrum knock-down of the tree canopy. The lower bound prediction that, for an order-of-magnitude decrease in body length, there should be a 560-fold increase in the number of individuals, is indicated by the lower dashed line on each graph; the upper bound prediction—a 1,780-fold increase for an order-of-magnitude decrease in body length—is shown by the upper dashed line. Regression lines were fitted through the log-transformed data, the anti-logarithm of the slopes of the lines were taken and are as follows: *a*, 390; *b*, 1,040 (●) and 700 (○); *c*, 360 (●) and 980 (○); *d*, 21,800 (●) and 530 (○); *e*, 1,440. The decline in the numbers of individuals of <1 mm body length may be due to a variety of factors¹², not least of which is the difficulty in sampling very small arthropods. Consequently, when fitting regression lines to the data, in *a-d* the 1 mm body length data point and in *e* the first three data points were omitted.



tions predict an increase of between $178 \times 3.16 = 560$ -fold and $178 \times 10 = 1,780$ -fold in the number of individual animals⁷.

Data on the number of individual animals of different body lengths collected from the surface of vegetation are scarce. Figure 2 shows such data for sweep-net samples of terrestrial arthropods¹⁵⁻¹⁷ and original data collected by pyrethrum knock-down of a tree (*Betula pubescens*) canopy. All the data show a steep increase in the number of individuals for an order-of-magnitude decrease in body length, in general agreement with our predictions (Fig. 2).

The above heuristic calculations may be reversed, allowing prediction of the expected fractal dimension of the surface of vegetation, given the mean of the fitted slopes for the distribution of individuals' body lengths shown in Fig. 2. This calculation results in a prediction of the fractal dimension of the surface of vegetation of $D = 2.78$.

Clearly, approximate agreement between data and our predictions does not prove that the fractal nature of vegetation contributes to a steep increase in number of individuals as arthropods get smaller. Moreover, different values for D (1.28–1.79; Table 1) and the metabolic rate exponent (0.62–0.86; refs 13,14) lead to different predicted slopes in Fig. 2, as will the precise form of the relationship between the fractal dimension of a surface and transects across that surface. It remains to be seen whether detailed differences in the fitted slopes apparent in Fig. 2 can be explained by differences in the fractal dimension of particular plant surfaces, and by the biology of associated arthropods.

We know of no other attempts to explain patterns in the body size distributions of arthropods living on plant surfaces, although previous studies have examined species/frequency distributions for animals of different sizes^{12,18}. Theoretical and empirical links between patterns in the number of species and number of individuals of different body lengths constitute an important unsolved problem.

D.R.M. gratefully acknowledges the financial support of the NERC. Colleagues at York and Professor B. Mandelbrot provided helpful criticism of the manuscript.

Received 15 October 1984; accepted 20 February 1985.

1. Mandelbrot, B. B. *The Fractal Geometry of Nature* (Freeman, New York, 1983).
2. Avnir, D., Farin, D. & Pfeifer, P. *Nature* **308**, 261–263 (1984).
3. Mandelbrot, B. B., Passoja, D. E. & Paullay, A. J. *Nature* **308**, 721–722 (1984).
4. Mandelbrot, B. B. *Science* **156**, 636–638 (1967).
5. Burroughs, P. A. *Nature* **294**, 240–242 (1981).
6. Bradbury, D. G., Reichelt, R. E. & Green, D. G. *Mar. Ecol. Prog. Ser.* **14**, 295–296 (1984).
7. Lawton, J. H. in *Plant Surfaces* (eds Juniper, B. E. & Southwood, T. R. E.) (Edward Arnold, London, in the press).
8. Richardson, L. F. *Gen. Syst.* **6**, 139–187 (1961).
9. Paumgartner, D., Losa, G. & Weibel, E. R. *J. Microsc.* **121**, 51–63 (1981).
10. Rigaut, J. P. *J. Microsc.* **131**, 41–54 (1984).
11. Orford, J. D. & Whalley, W. B. *Sedimentology* **30**, 655–668 (1983).
12. May, R. M. in *Diversity of Insect Faunas* (eds Mound, L. A. & Waloff, N.) 188–204 (Blackwell, Oxford, 1978).
13. Peters, R. H. *The Ecological Implications of Body Size* (Cambridge University Press, 1983).
14. Schmidt-Nielsen, K. *Scaling: Why is Animal Size so Important?* (Cambridge University Press, 1984).
15. Janzen, D. H. & Schoener, T. W. *Ecology* **49**, 96–110 (1968).
16. Janzen, D. H. *Ecology* **54**, 659–686 (1973).
17. Andrews, R. M. *Breviora* **454**, 1–51 (1979).
18. Hutchinson, G. E. & MacArthur, R. H. *Am. Nat.* **93**, 117–125 (1959).

Education policy and the heritability of educational attainment

A. C. Heath*, K. Berg†, L. J. Eaves*, M. H. Solaas†, L. A. Corey*, J. Sundet‡, P. Magnus† & W. E. Nance*

* Department of Human Genetics, Medical College of Virginia, Richmond, Virginia 23298, USA

† Institute of Medical Genetics, and ‡ Institute of Psychology, University of Oslo, PO Box 1036, Blindern, Oslo 3, Norway

Many workers assume that genetically determined differences in intellectual ability¹⁻⁵ will be influenced little by changes in educational policy or other environmental interventions^{6,7}. Others^{8,9}, however, have suggested that increasing equality of educational opportunity will lead to an increase in the heritability of educational attainment. The resolution of this issue has been delayed until now because of the extremely large sample sizes which would be required¹⁰. Education data on twins and their parents, from the Norwegian twin panel^{11,12}, provide a unique opportunity to determine the impact on the heritability of educational attainment of the more liberal social and educational policies introduced in Norway after the Second World War¹³. As reported here, for individuals born before 1940 there is a strong effect of family background on educational attainment, accounting for 47% of the variance, though genetic factors account for an additional 41% of the variance. For females born after 1940 and before 1961, the relative importance of genetic (38–45%) and familial environmental (41–50%) differences changes very little. For males born during the same period, the broad heritability of educational attainment has increased substantially (67–74%), and the environmental impact of family background has correspondingly decreased (8–10%). For males, at least, having well-educated parents no longer predicts educational success, as measured by duration of education, independent of the individual's own innate abilities.

Adult twins from the Norwegian twin panel, a population-based register of like-sexed twins born in Norway between 1915 and 1961 and presently resident there, were asked, by mailed questionnaire, to report the duration of their own education (I, 0–7 yr; II, 8–9 yr; III, 10–12 yr; IV, over 12 years of education) and that of both their parents. Questionnaires were returned by one or both twins from 8,389 families. Each member of 4,608 twin pairs reported his own educational level. Although well-educated twin pairs were overrepresented in the sample of respondents, the sample obtained was more nearly representative of the adult population than the samples obtained in most volunteer studies¹¹. Polychoric correlations¹⁴ for educational attainment were computed between male and female monozygotic (MZ) and dizygotic (DZ) twin pairs, between the parents of twins, and between parent and offspring for twins born before 1940, during 1940–49 and during or after 1950 (Table 1). (We report correlations rather than variances and covariances¹⁵, as we are not making any predictions about the effects of social policy on the variance in educational attainment.) Data from

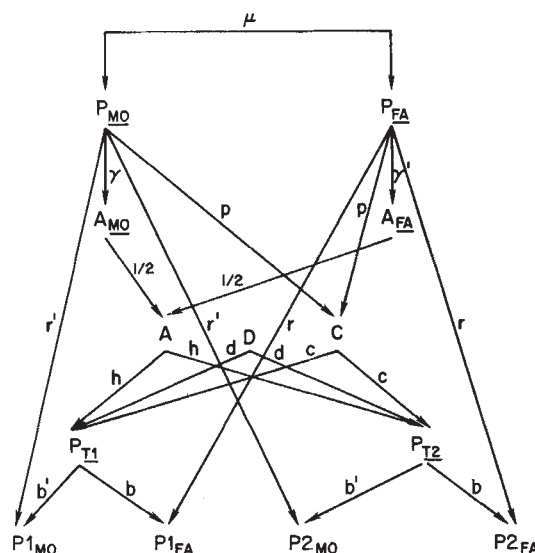


Fig. 1 Path diagram for MZ twins and their parents.

all the twins were used in computing the two-way contingency tables from which polychoric correlations between parent and offspring were calculated. Data for each member of a twin pair were treated as independent observations in computing the parent/offspring contingency tables. Data from only the first twin to report were used to compute the parent/parent contingency tables.

We found that correlations between male and female MZ twins have changed little over the time period studied, implying that the total contribution of familial factors (both genetic and environmental) to variation in educational attainment has remained constant. The change in the correlation between parents (0.85 to 0.79), though statistically significant because of the very large sample sizes, is also very small. There is a slight decline in the correlation between female DZ twins for twins born after 1940 (0.75 to 0.68), but a much more pronounced decline for male DZ twins (0.77 to 0.48); such a decline, relative to the MZ correlation, would be predicted if the influence of genetic differences on educational attainment has become more important in twins born after 1939. A corresponding decline in the parent/offspring correlations also occurs, the change in the parent/son correlations being greater than the change in the parent/daughter correlations.

Combining self-report data from the twins with the education data which they report for their parents allows us to overcome many of the limitations of the classical twin design¹⁵⁻¹⁷. Provided that the effects of dominance or genotype $\times$ age interaction are not too great, this design will resolve the contributions to variation in educational attainment of additive gene action, the environmental effects of parental education, other environmental effects shared by twins or siblings which are unrelated to parental education, and the genetic consequences of assortative mating. We shall not, however, be able to resolve alternative

Table 1 Secular changes in correlations between twins and their parents

	1915–39			Twins' year of birth 1940–49			1950–60		
	n	r	s.e.	n	r	s.e.	n	r	s.e.
Male MZ twins	259	0.86	0.03	253	0.82	0.03	370	0.85	0.02
Female MZ twins	405	0.89	0.02	342	0.85	0.03	518	0.89	0.02
Male DZ twins	313	0.77	0.04	284	0.48	0.06	463	0.47	0.04
Female DZ twins	425	0.75	0.03	400	0.68	0.04	576	0.66	0.03
Mother/son	1,652	0.74	0.02	1,456	0.61	0.03	1,809	0.51	0.02
Mother/daughter	2,001	0.72	0.02	1,694	0.57	0.03	2,179	0.55	0.02
Father/son	1,631	0.73	0.02	1,433	0.64	0.02	1,786	0.55	0.02
Father/daughter	1,971	0.74	0.02	1,660	0.79	0.02	2,121	0.60	0.02
Mother/father	2,519	0.85	0.01	2,027	0.79	0.02	2,542	0.79	0.01

Table 2 Correlations between educational levels of twins and educational levels reported by them for their parents

1915-39							Twin birth-cohort 1940-49						1950-60					
(1)	(2)	(3)	(4)	(5)	(6)		(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Male MZ twins (<i>n</i> = 193)							(1940-49 <i>n</i> = 203)						(1950-60 <i>n</i> = 255)					
(1)	1.00	0.03	0.07	-	0.06	0.07	1.00	0.03	0.06	0.07	0.06	0.07	1.00	0.03	0.06	0.07	0.06	0.06
(2)	0.86	1.00	0.09	-	0.07	0.06	0.83	1.00	0.07	0.07	0.06	0.06	0.86	1.00	0.07	0.07	0.06	0.06
(3)	0.74	0.60	1.00	-	0.05	0.06	0.72	0.63	1.00	0.03	0.05	0.07	0.52	0.53	1.00	0.01	0.03	0.04
(4)	0.73	0.70	0.95	1.00	-	-	0.61	0.67	0.91	1.00	0.05	0.05	0.48	0.50	0.96	1.00	0.04	0.04
(5)	0.75	0.64	0.87	0.84	1.00	0.02	0.71	0.65	0.81	0.78	1.00	0.01	0.58	0.56	0.82	0.76	1.00	0.01
(6)	0.69	0.71	0.83	0.79	0.93	1.00	0.57	0.66	0.71	0.82	0.96	1.00	0.54	0.56	0.78	0.80	0.97	1.00
Female MZ twins (<i>n</i> = 310)							(1940-49 <i>n</i> = 284)						(1950-60 <i>n</i> = 361)					
(1)	1.00	0.02	-	0.06	0.06	0.06	1.00	0.03	0.06	0.07	0.05	0.05	1.00	0.02	0.05	0.06	0.05	0.05
(2)	0.90	1.00	-	0.06	0.06	0.06	0.84	1.00	0.07	0.06	0.07	0.06	0.89	1.00	0.05	0.05	0.04	0.05
(3)	0.80	0.71	1.00	-	-	-	0.63	0.49	1.00	-	0.04	0.06	0.55	0.58	1.00	0.01	0.03	0.03
(4)	0.74	0.72	0.98	1.00	0.04	0.04	0.61	0.64	0.95	1.00	0.05	0.05	0.50	0.58	0.97	1.00	0.03	0.03
(5)	0.71	0.65	0.88	0.85	1.00	0.01	0.68	0.48	0.83	0.72	1.00	-	0.61	0.65	0.83	0.77	1.00	0.01
(6)	0.67	0.67	0.79	0.86	0.96	1.00	0.70	0.63	0.71	0.78	0.96	1.00	0.56	0.65	0.77	0.82	0.95	1.00
Male DZ twins (<i>n</i> = 243)							(1940-49 <i>n</i> = 235)						(1950-60 <i>n</i> = 318)					
(1)	1.00	0.04	0.05	0.06	0.04	0.05	1.00	0.06	0.07	0.08	0.06	0.07	1.00	0.05	0.06	0.06	0.05	0.06
(2)	0.78	1.00	0.06	0.05	0.05	0.04	0.47	1.00	0.07	0.07	0.07	0.06	0.46	1.00	0.06	0.06	0.06	0.06
(3)	0.78	0.73	1.00	0.02	0.02	0.04	0.55	0.56	1.00	0.03	0.04	0.05	0.49	0.40	1.00	0.03	0.04	0.05
(4)	0.71	0.76	0.96	1.00	0.04	0.04	0.49	0.57	0.90	1.00	0.06	0.05	0.40	0.57	0.86	1.00	0.05	0.04
(5)	0.79	0.73	0.92	0.84	1.00	0.02	0.63	0.54	0.85	0.71	1.00	0.02	0.54	0.41	0.73	0.65	1.00	0.02
(6)	0.73	0.79	0.85	0.84	0.95	1.00	0.45	0.60	0.74	0.74	0.91	1.00	0.44	0.52	0.61	0.76	0.91	1.00
Female DZ twins (<i>n</i> = 309)							(1940-49 <i>n</i> = 319)						(1950-60 <i>n</i> = 406)					
(1)	1.00	0.04	0.05	0.06	0.05	0.05	1.00	0.04	0.07	0.07	0.06	0.06	1.00	0.04	0.05	0.05	0.04	0.04
(2)	0.78	1.00	0.05	0.05	0.05	0.04	0.72	1.00	0.06	0.06	0.05	0.05	0.69	1.00	0.05	0.05	0.05	0.05
(3)	0.75	0.75	1.00	-	0.03	0.05	0.49	0.60	1.00	-	0.05	0.05	0.56	0.46	1.00	0.01	0.03	0.03
(4)	0.65	0.76	0.96	1.00	0.05	0.05	0.39	0.62	0.95	1.00	0.05	0.05	0.48	0.52	0.94	1.00	0.04	0.03
(5)	0.74	0.77	0.89	0.80	1.00	0.01	0.57	0.64	0.76	0.69	1.00	0.01	0.60	0.49	0.83	0.70	1.00	0.01
(6)	0.73	0.80	0.81	0.77	0.96	1.00	0.47	0.65	0.72	0.75	0.96	1.00	0.60	0.56	0.78	0.81	0.91	1.00

Polychoric correlations are given in the lower triangle of each matrix, their standard errors in the upper triangle. (1) First twin's education, (2) second twin's education, (3) mother's education (first twin's report), (4) mother's education (second twin's report), (5) father's education (first twin's report), (6) father's education (second twin's report). -, No estimate of standard error possible due to very low expected cell frequencies.

hypotheses about mechanisms of mate selection and environmental transmission within families¹⁸. As we have separate reports of parental educational levels by each member of a twin pair, we can check the validity of such data.

Table 2 shows matrices of polychoric correlations, broken down by twin group and birth-cohort, between the educational levels of each twin, and the educational level reported by each twin for each parent; only twin pairs where each twin reported the educational levels of himself and both parents were used in computing these correlations. For some of the polychoric correlations in Table 2 no standard errors are given, because of very low expected frequencies for one or more cells of the two-way 4 × 4 tables from which the polychoric correlations are derived. Such coefficients may be strongly biased estimates of the true population correlations¹⁴ and so were not used in any data analyses. Agreement between twins is very good, correlations between educational levels reported by each twin for the same parent ranging from 0.86 to 0.98. However, the correlation between a twin's educational level and that which he reports for either parent is consistently higher than the correlation between his own educational level and the educational level reported by his co-twin for the same parent. Fortunately, we can allow for the effects of this bias, by model-fitting.

Figure 1 presents a path model^{12,19} for deriving expected correlations between the educational attainments of MZ twins and the educational attainments reported by them for their parents: P denotes true educational attainment (we assume that the twins make no error in reporting their own educational level), and P1 and P2 the educational attainments reported by first and second twins, respectively, for mother (subscript MO) or father (subscript FA). The standardized path regressions of true phenotypic value on additive genetic deviation (A), dominance deviation (D), familial environmental deviation (C) and random environmental deviation (E: omitted from the diagram to simplify presentation) are denoted by parameters *h*, *d*, *c* and

e, or for females, in models which allow for sex-dependent effects, *h'*, *d'*, *c'* and *e'*. The path regression of additive genetic deviation on true phenotypic value is denoted by the parameter γ' in males, or γ in females, where $\gamma' = h + ac$, $\gamma = h' + ac'$, and *a* is the genotype/environmental correlation which may be expressed as a function of the other parameters of the model³. The path regressions of reported parental educational attainment on true parental educational attainment and on the twin's educational attainment are denoted by *b* and *r* for fathers, *b'* and *r'* for mothers. Path regressions of offspring familial environmental value on true maternal or paternal educational attainment are denoted by the environmental transmission parameter *p*. The parameter μ denotes the correlation between the educational attainments of spouses, which is assumed to arise through phenotypic assortative mating²⁰. An equivalent diagram can be drawn for DZ twins, and expected correlations derived.

Models were fitted to the polychoric correlations in Table 2 by nonlinear weighted least-squares, using the standard errors of the polychoric correlations as weights, and treating the correlations as though they were independent. Strictly, these correlations are not independent. However, ignoring the correlation between correlations has been found to have little effect on the parameter estimates obtained²¹. Model-fitting yields estimates of model parameters and provides an approximate χ^2 value with which to assess the fit of a given model, and compare its fit with that of other models. The validity of the χ^2 statistic will depend on how closely the sampling distribution of the polychoric correlation coefficients approximates to the multivariate normal distribution. With these very large sample sizes we expect the approximation to be close.

For twins born before 1940, there is no evidence for sex differences in the familial transmission of educational attainment. A purely environmental model gives a good fit to the data ($\chi^2_{24} = 42.18$, $P = 0.55$), although allowing for additive genetic variance gives a significant improvement in fit ($\chi^2_1 = 7.81$,

$P < 0.01$). Parameter estimates obtained for the best model are $h = 0.64$, $c = 0.53$, $\mu = 0.86$, $p = 0.19$, $r = 0.84$ and $b = 0.20$, so that 41% of the variance in educational attainment is attributable to additive genetic factors (h^2), 28% to the environmental effects of family background (c^2), and 19% to genotype/environmental covariation ($2hac$), the remaining 13% of the variance being attributable to environmental effects not shared by relatives. Our estimate of the importance of genotype/environmental covariation depends critically on the assumptions made about the mechanisms of environmental transmission and mate selection, which cannot be tested with these data. If we were to make certain alternative assumptions, no genotype/environmental covariation would be predicted⁵ and the contribution of the family environment would be estimated at 47%.

For twins born after 1939, we find significant heterogeneity of gene expression and environmental effects across sexes. Purely environmental models are rejected with a high level of significance (1940–49: $\chi^2_{50} = 89.44$, $P < 0.001$; 1950–60: $\chi^2_{53} = 169.55$, $P < 0.001$). Models which allow for both additive gene action and environmental transmission from parent to offspring yield negative estimates for the environmental transmission parameter p , which suggests that any environmental effects of parents on their offspring are being masked by genetic dominance in this design. Models which allow for sex-dependent additive gene action, dominance and environmental effects shared by twins, but no environmental transmission from parent to offspring, give the best fits to these data (1940–49: $\chi^2_{48} = 52.61$, d.f. = 48; $\chi^2_{51} = 56.59$, d.f. = 51). Parameter estimates obtained for this model are, for twins born between 1940 and 1949: $h = 0.70$, $h' = 0.67$, $d = 0.50$, $d' = 0$, $c = 0.29$, $c' = 0.64$, $\mu = 0.72$, $r = 0.82$, $b = 0.31$; and, for twins born between 1950 and 1960: $h = 0.52$, $h' = 0.53$, $d = 0.63$, $d' = 0.32$, $c = 0.45$, $c' = 0.71$, $\mu = 0.73$, $r = 0.82$, $b = 0.40$. These estimates imply that genetic factors (including genetic dominance) account for 74% of the variance in educational attainment in male twins born in 1940–49, and 67% in male twins born in 1950–60, but account for only 45% of the variance in female twins born in 1940–49, and 38% in female twins born after 1949. Conversely, environmental effects shared by twins account for only 8% (1940–49) or 20% (1950–60) of the variance in males, but for 41% (1940–49) or 50% (1950–60) in females. As genetic dominance is a major source of variation in IQ^{1–3,22}, the absence of evidence for dominance in the pre-1940 sample might be interpreted as evidence against a major influence of IQ on educational attainment in individuals born before 1940. However, dominance may be masked by strong environmental transmission in these data.

The results presented here are clearly consistent with the hypothesis that the importance of genetic influences on educational attainment is subject to secular change. Other explanations of our findings seem implausible. Perhaps more pre-war monozygotic male twin pairs have been misclassified as dizygotic, thus inflating the pre-war dizygotic male correlation? The proportion of pre-war male twins who are dizygotic (45.3%) differs little from the proportion of dizygotic twins in the other twin cohorts (44.4–48.8%). Perhaps a 'special twin environment' effect²³ is involved? As there has been a progressive decrease in the influence of parental educational levels on offspring educational attainment (see Table 1), we would expect the difference between male MZ and DZ twin correlations to have declined rather than increased in recent years. The most likely explanation, confirming the hypothesis of Scarr-Salapatek⁸, is that increased educational opportunity has led to an increased dependence of educational attainment on innate ability.

This work was supported in part by NIH grants GM30250, HL28922 and HD10291, and by grants from the Norwegian Research Council for Science and the Humanities.

Received 6 November 1984; accepted 8 February 1985.

1. Jinks, J. L. & Fulker, D. W. *Psychol. Bull.* **73**, 311–349 (1970).
2. Jinks, J. L. & Eaves, L. J. *Nature* **248**, 287–289 (1974).
3. Loehlin, J. C. *Behav. Genet.* **8**, 415–426 (1978).
4. Rice, J., Cloninger, C. R. & Reich, T. *Behav. Genet.* **10**, 73–92 (1980).
5. Rao, D. C. *et al. Genet. Res.* **39**, 187–198 (1982).

6. Jensen, A. R. *Genetics and Education* (Methuen, London, 1972).
7. Jensen, A. R. *Educability and Group Differences* (Methuen, London, 1973).
8. Scarr-Salapatek, S. *Science* **174**, 1223–1228 (1971).
9. Scarr-Salapatek, S. *Science* **174**, 1285–1295 (1971).
10. Eaves, L. J. & Jinks, J. L. *Nature* **240**, 84–88 (1972).
11. Magnus, P., Berg, K. & Nance, W. E. *Clin. Genet.* **24**, 103–112 (1983).
12. Heath, A. C. *et al. Behav. Genet.* (in the press).
13. *Educational Statistics* (Central Bureau of Statistics, Oslo, 1978).
14. Olsson, U. *Psychometrika* **44**, 443–460 (1979).
15. Eaves, L. J. *et al. Heredity* **41**, 249–320 (1978).
16. Eaves, L. J. *J. R. Statist. Soc. Ser. A* **140**, 324–355 (1976).
17. Martin, N. G. *et al. Heredity* **40**, 97–116 (1978).
18. Heath, A. C. & Eaves, L. J. *Behav. Genet.* **15**, 15–30 (1985).
19. Wright, S. *Ann. math. Statist.* **5**, 161–215 (1934).
20. Fisher, R. A. *Trans. R. Soc. Edinb.* **52**, 399–433 (1918).
21. Rao, D. C. *et al. Behav. Genet.* **7**, 147–159 (1977).
22. Bashi, J. *Nature* **266**, 440–442 (1977).
23. Kamin, L. J. *The Science and Politics of IQ* (Wiley, New York, 1974).

The genetic basis of Haldane's rule

Jerry A. Coyne

Department of Zoology, University of Maryland, College Park, Maryland 20742, USA

'Haldane's rule', formulated by J. B. S. Haldane in 1922, states that: "When in the F_1 offspring of two different animal races one sex is absent, rare, or sterile, that sex is the heterozygous [heterogametic] sex" (ref. 1). His rule is now known to apply in mammals, lepidopterans, birds, orthopterans and dipterans^{1–5}. In *Drosophila*, for example, Bock⁶ cites 142 cases of interspecific hybridizations that produce one sterile and one fertile sex in the offspring, all but one of these crosses yielding sterile XY males and fertile XX females. Despite much speculation, however, the genetic basis of Haldane's rule remains unknown. Haldane himself rejected the simple explanation that males are innately more sensitive than females to the effects of hybridization because groups with heterogametic females (such as birds and butterflies) usually show female sterility in hybrids, so that heterogamety itself is the critical feature. He and others^{1,2,4,7} suggested that heterogametic infertility or inviability in hybrids arises by a genetic imbalance between X chromosomes and autosomes. An alternative explanation^{5,8,9} is that this syndrome is caused by a mismatch of X and Y chromosomes. Here I show that in the *Drosophila melanogaster* subgroup, Haldane's rule for fertility apparently arises from a genetic interaction between X and Y chromosomes and not from an imbalance between sex chromosomes and autosomes. This finding has important implications for understanding the evolution of interspecific reproductive isolation.

Haldane's original explanation for asymmetrical sterility or viability rests on the fact that homogametic hybrids have a complete haploid set of chromosomes from each parent, but heterogametic hybrids lack an X chromosome from one of the parental species. The latter hybrids may lack the gene products from both autosomes and sex chromosomes of a complete haploid set which are thought to be necessary for hybrid fertility or viability. The genetic implications of this model are that epistatic interactions between conspecific X chromosomes and autosomes are necessary for proper gene action^{7,10}. Alternatively, different species may differ in their linkage relationships; some loci present on the X chromosome of one might be carried on the autosomes of the other. Homogametic hybrids would then have the normal dose of each locus, while heterogametic hybrids would suffer from duplication and deficiencies¹¹.

One way to test this X-autosome imbalance theory is to produce interspecific hybrid females having both X chromosomes from a single parental species. Because such females will have the same degree of X-autosome mismatch as that of sterile males, they would be expected to be sterile. Only in three species of *Drosophila* (*D. simulans*, *D. mauritiana* and *D. sechellia*) is it possible to carry out such a test, as the three species will hybridize to produce fertile females and sterile males^{12–15} and there exists in *D. simulans* a stock with attached X chromosomes. *D. simulans* is cosmopolitan, and *D. mauritiana* and *D. sechellia*

are found on the island of Mauritius and on Cousin and Praslin (Seychelles), respectively. The two island species probably arose by independent colonization of the islands by a *simulans*-like ancestor^{13,14}.

I used the attached-X *D. simulans* stock to test the fertility of interspecific female hybrids having both of their X chromosomes from *D. simulans* but two species-specific sets of autosomes, one from *D. simulans* and one from another species. The sector of Fig. 1 enclosed by dotted lines shows the cross producing such hybrids. Virgin hybrid females and control intraspecific females were tested for fertility by mass mating them for 72 h to males from both parental species, and then placing individual females in vials containing males from both parental species. Vials were inspected for female survival after 48 h, and for larvae and pupae after 1 week.

Table 1 shows that female hybrids with two *D. simulans* X chromosomes suffer no loss of fertility compared with the con-

Table 1 Fertility of attached-X females in interspecific and control crosses

Cross	Fertile	Non-fertile	Dead	Total
Control (<i>D. simulans</i>)	171	8	34	213
Interspecific hybrids				
<i>simulans</i> × <i>sechellia</i>	193	5	5	203
<i>simulans</i> × <i>mauritiana</i> (Bowling Green)	137	0	21	158
<i>simulans</i> × <i>mauritiana</i> (mixed stock)	124	2	19	145

For a description of the crosses, see text and Fig. 1 (cross enclosed by dotted lines).

trols, as the control cross shows the highest proportion of sterile females (0.044) and the three interspecific crosses are not heterogeneous for fertility ($G_2 = 5.32$, $P > 0.05$). The control cross also has the highest proportion of attached-X females not surviving the 48-h egg-laying period. Female hybrids therefore suffer no decrease in fertility when both of their X chromosomes come from a single species, and the X-autosome imbalance theory cannot be a correct explanation of hybrid sterility in this group.

An alternative (and largely ignored) explanation for Haldane's rule is that fertility is only possible when both X and Y chromosomes come from the same species^{5,8,9}. In *Drosophila*, nearly all loci on the Y chromosome affect spermatogenesis, and genetic interactions affecting spermatogenesis have been found between homologous loci on the X and Y chromosomes^{16,17}. This suggests that male fertility in *Drosophila* may depend strongly on interactions between the X and Y chromosomes. I investigated the validity of the X/Y interaction hypothesis by constructing two genotypes differing (as far as possible) only in the species identity of the Y chromosome. If 'deleterious' interactions between X and Y chromosomes from different species are the cause of Haldane's rule, then males with heterospecific sex chromosomes should be sterile.

Figure 1 shows the two crosses yielding males with different species combinations of X and Y chromosomes. Cross *a* produces males with a complete *D. simulans* X chromosome, a *D. mauritiana* Y chromosome, hybrid cytoplasm and an average of three-quarters of the autosomes from *D. simulans*. Cross *b* produces males with hybrid cytoplasm and with autosomes identical to those of cross *a*, but with a Y chromosome from *D. simulans* and a mixture of X chromosomal material from both species. A complicating variable in this comparison is that male *mauritiana/simulans* hybrids bearing the base of the *D. mauritiana* X chromosome in cross *b* produce no motile sperm¹⁸. The reason for this sterility is not known, but segregation of this region in this backcross leads to the expectation of sterility in at least half the male progeny. If X/Y chromosomal mismatch has no effect on sterility, males from cross *b* should have at most half the fertility of males from cross *a*. However, if X/Y mismatch makes a large contribution to sterility, male fertility

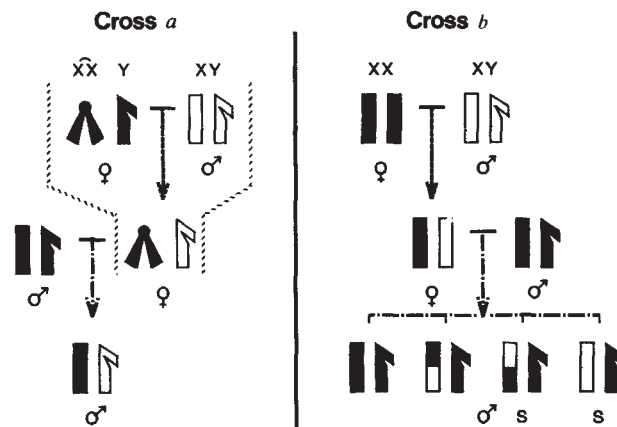


Fig. 1 Test of autosome-sex chromosome 'matching' hypothesis (cross enclosed by dotted lines): Attached-X *D. simulans* females [$C(1) y w, Y$], with sex chromosomes shown in black, are crossed to *D. mauritiana* or *D. sechellia* males, with sex chromosomes shown in white. Two strains of *D. mauritiana* were used, one from the Bowling Green Stock Center and the other a 'mixed' line of six isofemale strains provided by J. David. The resulting female interspecific hybrids have a haploid complement of autosomes from each species but both X chromosomes from *D. simulans*. They also contain a Y chromosome from their father, but this chromosome has no detectable phenotypic effects in such females and in particular does not affect fertility^{19,20}. Thus, this cross is not a test of the X/Y interaction theory. A control cross to assess the degree of intraspecific sterility was made between the attached-X *D. simulans* females and males from a conspecific strain collected in Oxnard, California. Test of the X-Y 'interaction' hypothesis (complete figure): Two genotypes were produced by backcrossing *D. simulans/D. mauritiana* hybrid females to *D. simulans* males from the Oxnard strain. One genotype (cross *a*) used as the original *D. simulans* female parent the attached-X stock of *D. simulans* (sex chromosomes in black); this was crossed to *D. mauritiana* males (sex chromosomes in white). The resulting F_1 hybrid females, which have free recombination among the autosomes²¹ and possess both X chromosomes from *D. simulans* and a Y chromosome from *D. mauritiana*, were backcrossed to *D. simulans* males. The resulting male offspring (shown at bottom left) have hybrid cytoplasm, a *D. simulans* X chromosome, a *D. mauritiana* Y chromosome and an average of three-quarters of the autosomes from *D. simulans*. The second cross (cross *b*) used as the original *D. simulans* female parent the wild-type Oxnard strain (sex chromosomes in black); these were crossed to *D. mauritiana* males (sex chromosomes in white). The F_1 hybrid females, which show free recombination, were backcrossed to *D. simulans* males, giving male offspring shown at bottom right. These males have hybrid cytoplasm, half of their X chromosomal material from *D. simulans*, half from *D. mauritiana* and a Y chromosome from *D. simulans* (various mixtures of X chromosomes are shown in the figure). The autosomes are identical in constitution to males from cross *a*. Genotypes with the base of the *D. mauritiana* X chromosome (shown by 'S') are known to have no motile sperm¹⁸.

in cross *a* should be lessened, and might perhaps be lower than that from cross *b*. Male fertility was measured as the proportion of 4-day-old virgin males with motile sperm¹⁸.

The results of these crosses are unambiguous. Males from cross *a* (with completely discordant X and Y chromosomes) had motile sperm in 5 of 615 individuals tested. The cross *b* males, however, who had a mixture of *D. simulans* and *D. mauritiana* X-chromosomal material and a *D. simulans* Y chromosome, showed motile sperm in 65 of 682 individuals tested, a value close to that obtained by other workers¹² for fertility in this backcross, and reflecting the segregation of at least five loci affecting sperm motility¹⁸. The difference in fertility between the crosses is highly significant ($G_1 = 57.7$, $P < 0.001$) and is in the direction predicted by the X/Y interaction theory. Given that at least half of the males from cross *b* are known from the outset to lack motile sperm, the substitution of a Y chromosome from *D. mauritiana* for one from *D. simulans* must reduce by at least 96% the incidence of individuals with motile sperm.

These two experiments, then, militate against the X-autosome imbalance theory for Haldane's rule and support the hypothesis that interactions between X and Y chromosomes from different species are largely responsible for heterogametic sterility in species hybrids. An example of such interactions may be the recent demonstration that the *Stellate* locus, which affects spermatogenesis in *D. melanogaster*, is found in multiple copies on both the X and Y chromosomes, and that Y-linked copies may regulate the transcription of those on the X. The sibling species *D. simulans*, however, lacks the *Stellate* gene on its X chromosome and has a reduced number of copies on the Y¹⁷.

Other possible explanations for hybrid male sterility are not supported by the crossing relationships of these species. For example, theories of deleterious interspecific interactions between the cytoplasm and X or Y chromosomes are not supported by the sterility of males from both reciprocal crosses between *D. mauritiana* and *D. simulans*¹². A theory of Y-autosome imbalance is not supported by the fact that genotypes

unbalanced in this way can have motile sperm in up to 20% of the individuals, while no individual with heterospecific X and Y chromosomes in the same cross has motile sperm¹⁸.

If X/Y interactions as a cause of heterogametic sterility prove to be widespread, theories of speciation must explain why X and Y chromosomes diverge genetically to produce reproductive barriers between geographically isolated populations. Sterility of the heterogametic sex often precedes all other reproductive and developmental disharmonies in hybrids, and as a first symptom of speciation deserves more attention from evolutionary biologists.

I thank J. David, V. Gvozdev, H. Robertson and I. Zhimulev for fly stocks, J. Beecham for assistance, and Ian Bock for permission to cite his unpublished results. G. Borgia, M. M. Green, L. Partridge, D. Reznick and J. Werren provided valuable criticism. This work has been supported by NSF grant BSR-83-18558 and grant 32221 from the Institute of General Medical Science of the NIH.

Received 28 September; accepted 11 December 1984.

1. Haldane, J. B. S. *J. Genet.* **12**, 101-109 (1922).
2. White, M. J. D. *Animal Cytology and Evolution* (Cambridge University Press, 1973).
3. White, M. J. D., Contreras, M., Cheney, J. & Webb, G. C. *Chromosoma* **61**, 127-148 (1977).
4. Harvey, A. W. *Biol. J. Linn. Soc.* **12**, 349-355 (1979).
5. Curtis, C. F., Langley, P. A. & Trewern, M. A. *Heredity* **45**, 405-410 (1980).
6. Bock, I. R. *Evol. Biol.* **18** (in the press).
7. Bacci, G. *Sex Determination* (Pergamon, Oxford, 1965).
8. Haldane, J. B. S. *The Causes of Evolution* (Longmans, Green, New York, 1932).
9. Fraccaro, M., Trepolo, L., Laudani, V., Marchi, A. & Jayakar, S. D. *Nature* **265**, 327-328 (1977).
10. Dobzhansky, T., Ayala, F. J., Stebbins, G. L. & Valentine, J. W. *Evolution* (Freeman, San Francisco, 1977).

11. Dobzhansky, T. *J. Genet.* **34**, 135-151 (1937).
12. David, J., Lemeunier, F., Tsacas, L. & Bocquet, C. *Ann. Genet.* **17**, 235-241 (1974).
13. Tsacas, L. & David, J. *Bull. Soc. Ent. Fr.* **79**, 42-46 (1974).
14. Tsacas, L. & Bächli, G. *Revue fr. Ent. (nouv. Ser.)* **3**, 146-150 (1981).
15. Lemeunier, F. & Ashburner, M. *Chromosoma* **89**, 343-351 (1984).
16. Williamson, J. H. in *The Genetics and Biology of Drosophila*, Vol. 1b (eds Ashburner, M. & Novitski, E.) 667-699 (Academic, London, 1976).
17. Livak, K. J. *Genetics* **107**, 611-634 (1984).
18. Coyne, J. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4444-4447 (1984).
19. Bridges, C. B. *Genetics* **1**, 1-52 (1916).
20. Hardy, R. W. *et al. Genetics* **107**, 591-610 (1984).
21. Sanchez, L. *Experientia* **38**, 448-449 (1984).

Genetic mapping and diagnosis of haemophilia A achieved through a *BclI* polymorphism in the factor VIII gene

Jane Gitschier, Dennis Drayna*, Edward G. D. Tuddenham†, Ray L. White* & Richard M. Lawn

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

* Howard Hughes Medical Institute and Department of Cellular, Viral, and Molecular Biology, University of Utah Medical Center, Salt Lake City, Utah 84132, USA

† Haemophilia Centre, Academic Department of Haematology, Royal Free Hospital, London WC1N 1BP, UK

Haemophilia A is the most common inherited bleeding disorder in man, affecting approximately 1 male in 10,000 (ref. 1). The disease is caused by a deficiency in the gene for factor VIII, a component of the intrinsic coagulation pathway. Due to the broad range of clotting activity in normal and heterozygous females, it is often difficult to confirm the status of women at risk for carrying the disease². A genetic marker in the form of a restriction fragment length polymorphism (RFLP) within or tightly linked to the factor VIII gene would serve as a tag for the haemophilia gene, thus allowing both accurate carrier detection and improved, earlier prenatal diagnosis by chorionic villi sampling^{3,4}. The recent isolation of the factor VIII gene^{5,6} has allowed a search for RFLPs within the gene, and we report here the identification of a common polymorphism within the factor VIII gene, revealed by the restriction enzyme *BclI*, which can be used diagnostically in about 42% of all families. Although the disease haemophilia A has been mapped to the distal portion of Xq (ref. 7), the *BclI* RFLP makes possible higher-resolution genetic linkage mapping with respect to other polymorphic markers on this portion of the X chromosome. We have established close linkage of the factor VIII gene to several useful RFLP markers, including the highly informative marker St14 (ref. 8). These markers should also be useful for prenatal diagnosis of haemophilia A and for detection of its carriers.

Table 1 Summary of the search for polymorphisms in the factor VIII gene

Enzyme	No. of bands	No. of chromosomes	Enzyme	No. of bands	No. of chromosomes
<i>AhaIII</i>	8	12	<i>MspI</i>	7	22
<i>ApaI</i>	8	12	<i>NciI</i>	5	6
<i>AvaII</i>	14	6	<i>NcoI</i>	9	9
<i>BamHI</i>	8	34	<i>NsiI</i>	14	11
<i>BanII</i>	7	12	<i>PstI</i>	13	12
<i>BclI</i>	12	12	<i>PvuII</i>	11	12
<i>BglI</i>	6	10	<i>RsaI</i>	14	12
<i>BglII</i>	11	36	<i>Sau3AI</i>	7	16
<i>BstEII</i>	6	12	<i>Sau96I</i>	13	10
<i>BstXI</i>	12	12	<i>ScaI</i>	13	10
<i>EcoRI</i>	15	45	<i>ScrFI</i>	10	10
<i>EcoRII</i>	15	6	<i>SphI</i>	10	12
<i>EcoRV</i>	8	12	<i>SstI</i>	10	20
<i>HaeIII</i>	7	6	<i>StuI</i>	9	12
<i>HincII</i>	9	12	<i>TaqI</i>	11	12
<i>HindIII</i>	9	18	<i>Tth111I</i>	3	12
<i>HinfI</i>	3	6	<i>XbaI</i>	5	12
<i>HpaI</i>	9	12	<i>XmnI</i>	10	10
<i>KpnI</i>	7	10			

Southern blots of human DNA (5 µg per lane) digested with each of the enzymes listed were hybridized to a full-length (9.0-kb) probe⁵ which included the complete factor VIII coding sequences and 3'-untranslated region. Male DNA, having one X chromosome, was generally used to eliminate ambiguity in the analysis of the band patterns. DNA samples were derived mainly from the blood of haemophiliacs at the Royal Free Hospital, London, although some samples were contributed by Utah families and employees of Genentech. The table lists the number of hybridizing bands and the number of independent chromosomes screened for each enzyme. Of these enzymes tested, only *BclI* revealed a factor VIII RFLP.

DNA samples from numerous individuals were examined for the presence of polymorphisms in the factor VIII gene. For each of 37 restriction enzymes, DNA samples were analysed by Southern blot⁹ hybridization to a full-length factor VIII DNA probe⁵. Table 1 gives, for each enzyme, the number of hybridizing bands and the number of chromosomes tested. For the enzymes with which the factor VIII gene has been completely mapped⁶ (*BamHI*, *EcoRI*, *KpnI*, *SstI*, *TaqI*), almost every exon

Fig. 1 Position of the *Bcl*I polymorphism within the factor VIII gene and an example of its usefulness in the diagnosis of haemophilia.

a, The factor VIII gene, containing 26 exons, is shown schematically⁶. Below it is an expanded portion of the gene covering exons 17 and 18, and the positions of the *Bcl*I sites (B) in the gene are indicated. The presence or absence of the middle site (in the intron just 3' of exon 18 and marked by an asterisk) is responsible for the two alleles of 879 bp and 1,165 bp. The location of the 647-bp *Stu*I/*Sca*I genomic probe fragment is also shown. **b**, A Southern blot⁹ of *Bcl*I-digested DNA (5 µg per lane) from members of two different families affected by haemophilia was hybridized to the 647-bp *Stu*I/*Sca*I genomic probe. Family 20 consists of a haemophiliac son (H), whose brother (B) is not a haemophiliac, his heterozygous mother (M), a known obligate carrier of haemophilia, and father (F). Family 22 consists of a haemophiliac (H), his heterozygous mother (M) and his brother (B), who is also a haemophiliac. Blood samples from both patients and families were obtained from the Royal Free Hospital, London.

Methods. DNA was prepared from peripheral blood lymphocytes¹⁷. The probe was labelled with [α -³²P]dCTP by calf thymus DNA priming. Nitrocellulose blots were hybridized¹⁸ in 50% formamide at 42 °C overnight and washed at 60 °C in 0.2 × SSC, 0.1% SDS. Films were exposed for 3 days at -70 °C with one intensifying screen.

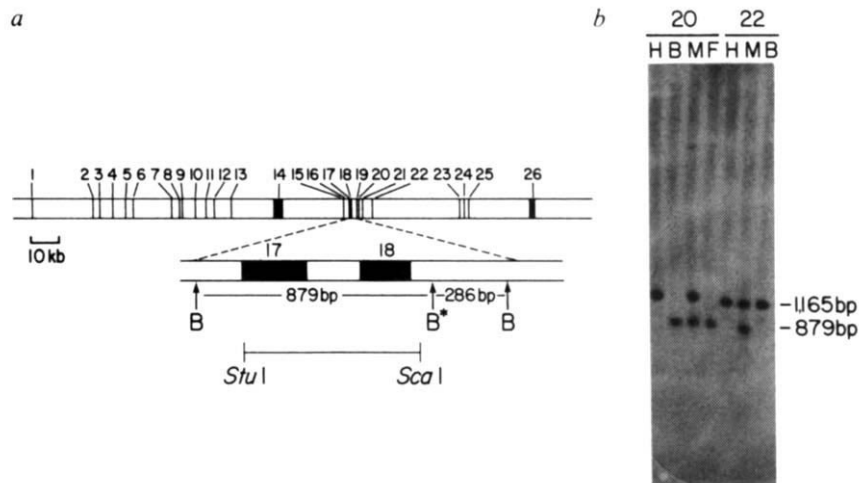
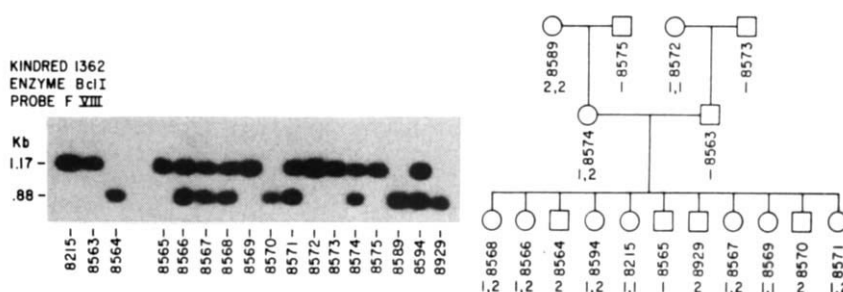


Fig. 2 An example of a three-generation Utah family showing the X-linked inheritance of the factor VIII *Bcl*I polymorphism. Normal human DNA from family members was isolated as described previously¹³, digested with *Bcl*I, electrophoresed on 0.8% agarose gel and blotted onto Magna 66 nylon membrane. The resulting blot was hybridized with the 647-bp *Stu*I/*Sca*I fragment of the factor VIII gene which had been ³²P-labelled. The numbers 1 and 2 represent the alleles of high and low relative molecular mass, respectively.



can be accounted for in the hybridizing bands, suggesting that exon-flanking restriction sites have been closely surveyed with many of these enzymes.

The complementary DNA probe revealed a polymorphism in the case of only one enzyme, *Bcl*I. Eleven hybridizing bands remained constant in all (male) individuals tested, whereas one faint band of either 0.9 or 1.1 kilobases (kb) appeared variably in different individuals. Genomic sequences⁶ of the exons and their flanking regions were examined to determine the location of the *Bcl*I polymorphism. Three *Bcl*I sites surrounding exons 17 and 18 were found to be in the configuration shown in Fig. 1a, suggesting that the polymorphic site is 3' of exon 18 and that the two-allele system is composed of 879- and 1,165-bp fragments.

For verification and clarity, a 647-bp *Stu*I/*Sca*I genomic probe (Fig. 1a) was prepared and hybridized to a Southern blot of haemophiliac DNAs. This probe reveals only the two variant bands. Of the 70 (male) haemophiliac DNA samples examined, 44 (63%) bear the 879-bp fragment and 26 (37%) the 1,165-bp fragment. Five of eight mothers of haemophiliacs were found to be heterozygous for the *Bcl*I polymorphism. Figure 1b shows an example of two families in which the haemophilia chromosome can be followed by its *Bcl*I allele. For those families in which the mother was heterozygous and in which the association between the *Bcl*I allele and the haemophilia gene was known, a prenatal diagnosis of haemophilia could be made. We have already used this polymorphism to predict the carrier status of one woman whose brother is a haemophiliac.

Our survey has revealed a paucity of RFLPs with the factor VIII probe. There are some indications that DNA polymorphisms may generally be more scarce in the human X chromosome

than in autosomes¹⁰. In the entire 9,000 base pairs (bp) of the messenger RNA coding region of the factor VIII gene sequenced from two individuals, we found only two nucleotide differences^{5,6}; this frequency is less than that found for the autosomal globin genes. In addition to the common *Bcl*I polymorphism, we have detected four different *Taq*I and one *Nsi*I rare variants. Each has been seen in one particular haemophiliac family, and two of the sequence variations may lead to premature termination of factor VIII translation¹¹. None of these restriction fragment length variants is common enough to be of general use.

The *Bcl*I polymorphism also serves as a genetic marker for mapping the factor VIII gene. Several useful RFLP markers on the distal portion of Xq have already been described and used for linkage mapping studies¹². These markers span the physical distance from Xq26 to Xq28 and cover a total genetic distance of 35 recombination units. Table 2 lists the RFLP markers used

Table 2 Markers used to map factor VIII

Marker	Polymorphism	Frequencies	Location	Ref.
St14	<i>Taq</i> I	0.36/0.20/0.15 0.12/0.11/0.04 0.005	Xq28	8
	<i>Msp</i> I	0.35/0.35/0.20 0.10		
DX13	<i>Bgl</i> II	0.65/0.35	Xq28	12
Factor IX	<i>Taq</i> I	0.65/0.35	Xq27	15
52A	<i>Taq</i> I	0.50/0.50	Xq27	12
<i>hprt</i>	<i>Bam</i> HI	0.77/0.16/0.07	Xq26	16

in this study, together with their physical and genetic locations. The linkage of these markers to the factor VIII *BclI* polymorphism was determined in a series of three-generation Utah families¹³; an example is shown in Fig. 2. The frequencies of the two *BclI* alleles in the Utah families we tested (totalling 63 X chromosomes) are 79% (879 bp) and 21% (1,165 bp) (Table 3). The combined data for the haemophilic and normal families are 71% (879 bp) and 29% (1165 bp) for 133 X chromosomes

Table 3 Linkage of factor VIII to other markers

Pairwise cross	No. recombinants/ total meioses	Recombination fraction	LOD
Factor VIII × St14	0/63	0	19.0
Factor VIII × DX13	0/27	0	7.8
Factor VIII × Factor IX	1/13	0.08	2.4
Factor VIII × 52A	6/27	0.22	1.92
Factor VIII × <i>hprt</i>	4/7	~0.50	—

examined. The predicted frequency of females heterozygous for this marker is therefore 42% (2 X .71 X .29).

No recombination was observed between the marker loci St14 and DX13 and the factor VIII locus. The lod score for linkage of factor VIII to St14 is 19 at a recombination fraction of 0.0, indicating that the factor VIII gene is within less than 4.6 recombination units (with 95% confidence) of this locus. This linkage analysis places factor VIII 8 recombination units distal to factor IX, 22 recombination units distal to 52A and ≥50 recombination units distal to the gene for hypoxanthine phosphoribosyl transferase. The linkage distances obtained with these other loci are consistent with the placement of factor VIII very near the loci St14 and DX13, and add confidence to this map assignment.

The tight linkage of the factor VIII gene to DX13 and St14 (one of the most polymorphic markers in the human) makes feasible carrier detection and prenatal diagnosis of the disease in those families in which the *BclI* polymorphism is not informative. These linkage results provide a high level of confidence in the close linkage of these markers to factor VIII. Since other studies showing linkage of these markers to the disease haemophilia A are based on fewer meioses^{8,14}, the added confidence derived from this study is important for genetic counselling of families affected by haemophilia A.

We thank Terri Goralka and Mark Hoff for their technical contributions, William Wood for helpful discussions, Alane Gray and Jeanne Arch for manuscript preparation, Marc Shuman (University of California, San Francisco) and Peter Jones (Royal Infirmary, Newcastle) for providing some of the haemophilic blood samples, and all patients and their families for their cooperation. D.D. is a fellow of the Muscular Dystrophy Association and R.L.W. is an investigator of the Howard Hughes Medical Institute. Part of this work was supported by Genentech, Inc.

Received 26 November 1984; accepted 24 January 1985.

- Galjaard, H. *Genetic Metabolic Disease* (Elsevier, New York, 1980).
- Graham, J. B. *Clin. Haemat.* **8**, 115-145 (1979).
- Mibashan, R. S. *et al. Lancet* **i**, 1309-1311 (1979).
- Williamson, R. *et al. Lancet* **ii**, 1125-1127 (1981).
- Wood, W. I. *et al. Nature* **312**, 330-337 (1984).
- Gitschier, J. *et al. Nature* **312**, 326-330 (1984).
- Fillipi, G. *et al. Am. J. hum. Genet.* **36**, 44-71 (1984).
- Oberle, I., Drayna, D., Camerino, G., White, R. & Mandel, J. L. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Southern, E. M. *J. molec. Biol.* **98**, 503-518 (1975).
- Cooper, D. N. & Schmidke, J. *Hum. Genet.* **66**, 1-16 (1984).
- Gitschier, J. *et al. Nature* (in the press).
- Drayna, D. *et al. Proc. natn. Acad. U.S.A.* **81**, 2836-2839 (1983).
- White, R. *et al. Nature* **313**, 101-105 (1985).
- Harper, K. *et al. Lancet* **ii**, 6-8 (1984).
- Camerino, G. *et al. Proc. natn. Acad. U.S.A.* **81**, 488-502 (1984).
- Nussbaum, R. W., Crowder, W. E., Nyhan, W. L. & Caskey, C. T. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4035-4039 (1983).
- Bell, G., Karam, J. H. & Rutter, W. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5759-5763 (1981).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).

Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations

Michael Graham, Jerry M. Adams & Suzanne Cory

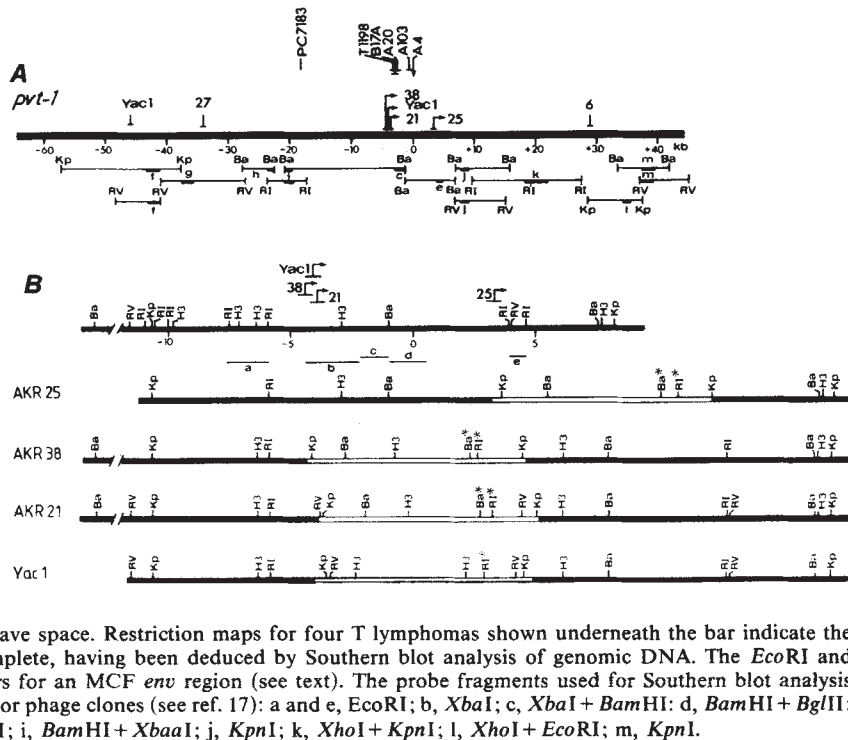
The Walter and Eliza Hall Institute of Medical Research,
PO Royal Melbourne Hospital, Victoria 3050, Australia

The frequent trisomy of murine chromosome 15 in T lymphomas suggests that it bears one or more genes conducive to T-cell neoplasia¹. One such gene seems to be *c-myc*, the oncogene frequently activated in B-lymphoid tumours either by retroviral insertion, as in the avian bursal lymphomas^{2,3}, or by a translocation to the immunoglobulin heavy-chain locus, as in the predominant t(12; 15) of murine plasmacytomas and the analogous t(14; 8) of human Burkitt lymphomas (for reviews, see refs 1, 4-6). The *c-myc* gene was strongly implicated in T-cell neoplasia when 15-25% of T lymphomas arising in AKR mice, a strain prone to leukaemia, were found to have retroviral inserts near *c-myc*^{7,8}. Proviruses near *c-myc* were also found in several T lymphomas induced by murine leukaemia viruses (MuLV) in both mice⁹ and rats¹⁰, but many of the rat thymomas bear an insert instead at one of several other common sites¹¹⁻¹³, at least two of which have murine homologues on chromosome 15 (refs 14, 15). We show here that some murine T lymphomas contain proviral inserts in the recently identified chromosome 15 locus for plasmacytoma variant (6; 15) translocations, which we have denoted *pvt-1* (refs 16, 17). Although 6; 15 breakpoints map cytogenetically to the same chromosome band as *c-myc*^{18,19}, the alterations of *pvt-1* in tumours occur at least 72 kilobases (kb) from the *c-myc* promoters¹⁷. The insertions in T lymphomas suggest that an altered *pvt-1* locus is conducive to neoplasia in T cells as well as B cells, possibly via long-range effects on *c-myc* expression.

Figure 1A shows the 108-kb region across *pvt-1* which we have cloned. In five t(6; 15)-bearing tumours (TEPC 1198, Baltnlm 17A, ABPC 20, ABPC 103 and ABPC 4), the breakpoint falls within a 4.5-kb region, while the site associated with the complex (6; 12; 15) rearrangement in plasmacytoma PC7183 (ref. 20) lies 14 kb farther to the left¹⁷. As several T lymphomas have a retroviral insert immediately 5' to *c-myc*⁷⁻¹⁰, close to the breakpoint region favoured by 12; 15 translocations^{21,22}, we reasoned that the *pvt-1* locus might also be a target for retroviral integration. We therefore looked for rearrangements in T-lymphoma DNAs by Southern blot analysis with a battery of probes from the cloned *pvt-1* DNA (*a-m* in Fig. 1). A judicious combination of digests and probes allowed us to survey 100 kb of the locus for alterations, as shown at the bottom of Fig. 1A. The panel of T lymphomas analysed included Yac-1, which was induced by Moloney murine leukaemia virus (M-MuLV)²³, and 32 primary tumours of AKR mice, a strain in which leukaemogenesis follows generation of recombinant endogenous retroviruses (MCF viruses) (see ref. 24 for a review). Rearrangements were detected in Yac-1 and five AKR tumours, as illustrated by representative blots in Fig. 2. One allele of *pvt-1* remained unaltered in all lines. The higher level of the rearranged fragment consistently noted in Yac-1 (see, for example, lanes 6 and 7 in Fig. 2) may reflect selective pressure for additional copies of an altered chromosome 15 at the expense of the normal one in some T-lymphoma lines^{7,25,26}.

Rearrangements in four of the six T lymphomas (Yac-1 and AKR 21, 38 and 25) map remarkably close to the 6; 15 translocation breakpoint cluster (Fig. 1A). The recombination site in each tumour was localized to the *KpnI* fragment spanning this region by showing that probes from either side (*a* and *e* in Fig. 1B) detected different rearranged *KpnI* fragments (illustrated in Fig. 2A for Yac-1 and AKR 25). The recombination point in Yac-1, AKR 38 and AKR 21 falls between regions *a* and *c*, as probe *c* hybridized to the same rearranged *KpnI* fragment as probe *e*

Fig. 1 The *pvt-1* locus on chromosome 15. **A**, The cloned 108-kb region across the *pvt-1* locus showing 6;15 breakpoints in six plasmacytomas (top) and insertions in six T lymphomas; arrows indicate the probable transcriptional orientation of the proviruses, where known. A 17-kb region around the ABPC 46;15 breakpoint is described in ref. 16; complete restriction maps of the clones defining the rest of the locus, the mapping of the four other plasmacytoma 6;15 breakpoints, and the location of the PC7183 breakpoint described by Van Ness *et al.*²⁰ within this locus are described elsewhere¹⁷. Below the bar, the digests and lettered unique-sequence probes (thicker lines) used to survey this region are indicated. An ~1-kb region at +28 kb on the map could not be surveyed with the available probes. Abbreviations for restriction endonuclease sites are: RI, *EcoRI*; Ba, *BamHI*; RV, *EcoRV*; H3, *HindIII*; and Kp, *KpnI*. **B**, Proviral insertions within the *pvt-1* locus. This expanded map of the region around the ABPC 4 breakpoint shows the location and orientation (arrows) of proviral inserts in T lymphomas; the bar indicates the possible error (± 0.3 kb) in placement of the insertion points and an 8-kb gap has been introduced to save space. Restriction maps for four T lymphomas shown underneath the bar indicate the proviral inserts (open bars); the maps are incomplete, having been deduced by Southern blot analysis of genomic DNA. The *EcoRI* and *BamHI* sites marked with an asterisk are markers for an MCF *env* region (see text). The probe fragments used for Southern blot analysis were derived from the following digests of cosmid or phage clones (see ref. 17): a and e, *EcoRI*; b, *XbaI*; c, *XbaI* + *BamHI*; d, *BamHI* + *BglII*; f, *BamHI* + *EcoRI*; g, *KpnI*; h, *HindIII* + *BamHI*; i, *BamHI* + *XbaI*; j, *KpnI*; k, *XhoI* + *KpnI*; l, *XhoI* + *EcoRI*; m, *KpnI*.



(not shown). The AKR 25 site, on the other hand, mapped between *c* and *e*. The AKR 25 rearrangement clearly resulted from an insertion, as probes from either side of the recombination site (*c* and *e*) hybridized to the same large rearranged *HindIII* fragment (shown for probe *c* in lane 9 of Fig. 2A); its size (~19.5 kb) compared with the germline fragment (10.6 kb) is consistent with insertion of ~9 kb of DNA lacking a *HindIII* site. Data obtained for the other three lines using flanking probes are consistent with the presence of a 9-kb insert bearing one *HindIII* site and one *EcoRI* site. Additional results from *BamHI*, *EcoRV* and *KpnI* digests allowed us to deduce the maps shown in Fig. 1B. The number of retroviral inserts in these tumours (see below) precluded using retroviral probes to verify that the inserts are proviruses, but all the data are consistent with that conclusion. The size of the inserts is that of a full-length (replication-competent) provirus, and the *KpnI* (and *EcoRV*) sites associated with LTRs allowed us to localize the inserts more precisely. The AKR 38, AKR 21 and Yac-1 inserts fall within a ~1-kb region located only ~1 kb from three of the 6;15 translocation breakpoints, while the AKR 25 insertion site lies 3 kb to the right of the ABPC 4 breakpoint (Fig. 1A). Thus, four retroviral inserts and five plasmacytoma breakpoints map within an 8-kb region.

Significantly, the four inserts have features characteristic of the MCF proviruses strongly implicated in leukaemogenesis²⁴. Like most MCF inserts found near *c-myc*⁷, all four bear an *EcoRI* site ~2 kb from one end and the three AKR inserts have a nearby *BamHI* site (Fig. 1B). These hallmarks of the *env* region (encoding the glycoprotein region gp70) of a leukaemogenic MCF virus are acquired by recombination with endogenous non-ecotropic proviruses^{24,27-30}. The hybrid gp70 is thought to promote leukaemogenesis by permitting retroviral invasion of thymocytes and/or providing them with a mitogenic stimulus²⁴. As the *env* gene lies downstream of the viral *gag* and *pol* genes, all four proviral inserts would be transcribed from left to right on the maps in Fig. 1.

The three other rearrangements lie farther from the plasmacytoma breakpoints (Fig. 1A); although we lack sufficient unique-sequence probes in the relevant flanking regions to characterize them extensively, it seems likely that they also reflect

proviral integration. The putative insert in AKR 6 maps ~29 kb to the right of the ABPC 4 translocation site (Fig. 1A), assuming that the acquired *KpnI* site (Fig. 2B) lies within an LTR, while that in AKR 27 maps ~16 kb to the left of the PC7183 translocation breakpoint. Curiously, Yac-1 appears to have an additional insert still farther from the PC7183 site (Fig. 1A). The apparent presence of two inserts within *pvt-1* in Yac-1 may reflect selection for further insertions during its long passage history.

Retroviral insertion within *pvt-1* is unlikely to be coincidental, as insertion is thought to be random³¹ and the 100-kb region surveyed here represents only 3.7×10^{-5} of the genome. Given ~10 inserts per AKR tumour (ref. 28 and our unpublished results), the chance of finding a single random insert within that

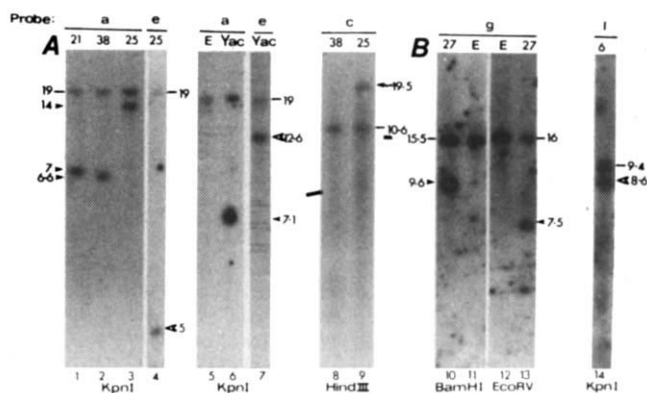


Fig. 2 Southern blots revealing *pvt-1* rearrangements in T lymphomas. **A**, Rearrangements in the region favoured for 6;15 translocations; **B**, more distant rearrangements. E denotes AKR embryonic DNA. Fragment sizes are in kilobases (kb). Bars indicate fragments from the unrearranged allele; solid arrowheads mark left-hand host-viral junction fragments, while open arrowheads indicate right-hand junction fragments. The *BamHI* fragment analysed in lanes 10 and 11 spans from ~-42 to -28 kb on the map in Fig. 1A. Letters denote probes described in Fig. 1 legend. The methods used for preparation of DNA of high relative molecular mass, its digestion by the indicated restriction endonucleases, subsequent electrophoretic fractionation, hybridization and autoradiography have been detailed elsewhere^{21,26}.

region in the 33 tumours analysed is $\sim 10 \times 33 \times 3.7 \times 10^{-5} = 0.01$, and the odds of finding seven inserts by chance are $\sim (10^{-2})^7 = \sim 10^{-14}$. The possibility that *pvt-1* is merely a preferred integration region, unrelated to oncogenesis, seems unlikely because chromosome 15 has not acquired endogenous proviruses³², nor is *pvt-1* merely a recombination 'hotspot', because translocations involving this band have not been associated with non-lymphoid tumours. Hence, it seems likely that the insertions contributed to T-lymphoma development. Whether *pvt-1* is related to any of the previously identified retroviral insertion regions¹¹⁻¹³ is not yet clear; it does not correspond to *pim-1*, a common locus of M-MuLV insertion in murine T lymphomas¹³, because their restriction maps differ markedly and *pim-1* does not map to chromosome 15 (A. Berns, personal communication). Conceivably, *pvt-1* may bear the murine homologue of one of the three common proviral insertion loci identified in rat thymomas (*mlvi-1*, *mlvi-2* and *RMoInt-1*)^{11,12}. Both *mlvi-1* and *mlvi-2* have homologues on mouse chromosome 15 (refs 14, 15), but that of *mlvi-2* does not appear to map to the distal region of chromosome 15 (ref. 14), where *myc* and *pvt-1* lie^{4,18,19}.

As the *pvt-1* breakpoints lie at least 72 kb from the *c-myc* promoters¹⁷, it is conceivable that *pvt-1* bears another oncogene, but we have found no evidence that it bears sequences related to known oncogenes¹⁷. In particular, the putative mammary oncogene *int-1* (ref. 33) and the *c-sis* gene, which both map to chromosome 15 (refs 33, 34), were not detected within the entire cloned *pvt-1* (or *c-myc*) regions; we found no rearrangements near *int-1* or *sis* when we screened *EcoRI* digests of the 32 AKR tumour DNAs with *int-1* probe *c* (see ref. 33), provided by Drs H. Varmus and T. Fung, and a *v-sis* probe.

To determine whether *pvt-1* is transcriptionally active, we used probes from the region near the major insertion/breakpoint region (from -11 kb to +4.8 kb on the map in Fig. 1B) to examine blots of poly(A)⁺ RNA from several T lymphomas and plasmacytomas; only low levels of transcripts were found and none was consistently elevated (or depressed) in lines having *pvt-1* alterations compared with those lacking any *pvt-1* rearrangement (ref. 17 and J.A., unpublished results). Thus we have not yet found an attractive candidate for a *pvt-1* gene that might be a target for the insertions and translocations.

Is retroviral insertion within the *pvt-1* locus associated with active *c-myc* expression, as are the variant (6; 15) plasmacytoma translocations involving *pvt-1* (refs 17, 19) and the variant Burkitt lymphoma translocations³⁵⁻³⁹? To address this question, we compared messenger RNAs from four T-lymphoma lines having *pvt-1* inserts, six AKR lines having inserts near *c-myc*⁷ and four AKR lines lacking detectable inserts within either locus. Nearly all the lines had comparable levels of the canonical ~2.4-kb *c-myc* mRNA (Table 1, Fig. 3). These levels fall within the range found in plasmacytomas bearing either the 6; 15 or 12; 15 translocation^{17,19,40-44} (illustrated respectively by ABPC4 and TEPC609 in Fig. 3) and are well above the range observed in normal lymphoid tissues or in most chemically induced T lymphomas examined previously⁴⁰. These results suggest that *c-myc* is activated in most retrovirally induced T lymphomas. In tumours with a provirus close to *myc*⁷⁻¹⁰, activation is probably mediated by the LTR enhancer, which frequently disrupts putative control sequences upstream from *myc*⁷. We speculate that insertions within *pvt-1*, and perhaps yet other loci, can also activate *c-myc*, either directly or indirectly (see below). Of the 35 retrovirally induced T lymphomas we have studied, 9 AKR tumours and 2 T lymphomas induced by Soule MuLV had insertions near *c-myc* (ref. 7 and Table 1), and 6 had insertions within *pvt-1* (Fig. 1A).

Our results suggest that alterations of *pvt-1* may be conducive to T-cell as well as B-cell neoplasia, perhaps by inducing constitutive *c-myc* expression. As cloned sequences from the *pvt-1* and *c-myc* loci establish that *pvt-1* alterations occur at least 72 kb from the *c-myc* promoters¹⁷, how might the effect be conveyed over such large distances? The 6; 15 translocations couple the left-hand side of the *pvt-1* locus to the *C_κ* enhancer^{16,17,20} and the insertions presumably introduce the LTR

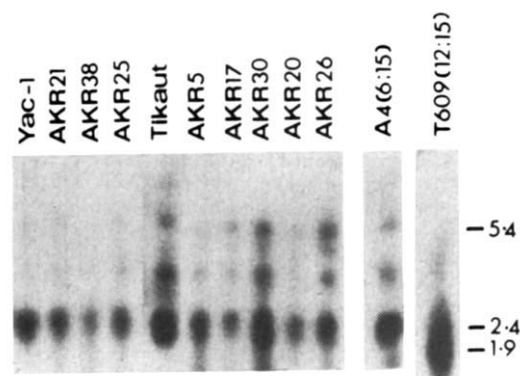


Fig. 3 Transcripts of the *c-myc* gene in 10 T lymphomas and 2 plasmacytomas. Blots were hybridized with a *myc* exon 3 probe (a *Sac*/*Xho* fragment)⁴¹. RNAs in the first four lanes derive from T lymphomas which have proviral inserts within *pvt-1*, while those in the second four lanes are from T lymphomas with inserts near *c-myc*⁷; AKR 20 and 26 have no inserts near *c-myc* or *pvt-1* (Table 1). Plasmacytoma ABPC 4 (A4) has a t(6; 15) and contains the standard ~2.4-kb transcripts^{17,19}, while TEPC609 has the typical *C_κ-myc* t(12; 15) fusion and yields only the 1.8–2.2-kb mRNAs derived from the rearranged gene⁴¹. The ~5.4-kb RNA is a nuclear precursor of the 2.4-kb mRNA⁴¹. Poly(A)⁺ RNA was prepared by a method⁴⁵ involving lysis of tumour tissue in SDS and proteinase K, followed by oligo(dT)-cellulose chromatography. RNA was denatured by heating for 10 min at 60 °C in formaldehyde, and 3 µg was fractionated on a 1.5% agarose gel in formaldehyde and transferred to nitrocellulose⁴⁶.

Table 1 T-lymphoma insertions and *c-myc* expression

Tumour	<i>c-myc</i> locus*	<i>pvt-1</i> locus†	<i>c-myc</i> mRNA‡
AKR 21	—	+	++
AKR 38	—	+	+
AKR 25	—	+	++
AKR 6	—	+	ND
AKR 27	—	+	ND
Yac-1	—	+	++
Tikaut	5'	—	+++
AKR 22	5'	—	++
AKR 12	5'	—	±
AKR 5	5'	—	+
AKR 16	5'	—	ND
AKR 17	5'	—	+
AKR 2	3'	—	ND
AKR 30	3'	—	++
AKR 14	3'	—	ND
AKR 20	—	—	++
AKR 19	—	—	++
AKR 26	—	—	++
AKR 10	—	—	+

Tumours denoted AKR arose spontaneously in mice 7–12 months old in our AKR mouse colony⁷. Tikaut is a cell line derived from an AKR tumour²⁷ and Yac-1 a cell line derived from a M-MuLV induced BALB/c tumour²². Tikaut is trisomic²⁵, but no cytogenetic data are available on the other lines.

* Insert detected within a 53-kb region across *c-myc* described elsewhere¹⁷, extending from 27 kb 5' to the *c-myc* promoters to 26 kb 3' of them. All 5' inserts map within 2 kb of the *c-myc* promoters⁷. The insert in AKR 30 maps ~24 kb 3' of the *c-myc* promoters⁷, and that in AKR 14 in the same vicinity (unpublished results), while that in AKR 2 appears to be in the 3'-untranslated portion of exon 3 (ref. 7).

† Insert detected within the 108-kb cloned region across *pvt-1* (see Fig. 1A).

‡ Estimate of the amount of *c-myc* mRNA bases on several blots like that in Fig. 3; +++ indicates the level of mRNA in Tikaut; ++, about half that level; +, about one-quarter; and ±, about one-eighth. ND, not determined.

enhancer (and promoter), but enhancers are generally thought to act over distances of only 5–10 kb. Conceivably, the translocations/insertions may alter a *pvt-1* gene that regulates *myc* expression, either disrupting a negative regulator or stimulating expression of a positive regulator, although we have not yet found a *pvt-1* transcript that is consistently depressed or stimulated by the *pvt-1* alterations. Alternatively, the effect may be conveyed along the chromosome (in *cis*) by disruption of long-range chromatin folding, although such effects

might not be expected from a 9-kb insert. For either model, it is tempting to think that *pvt-1* is involved in normal *myc* regulation.

We thank Jan Mitchell, Linda Byrt and Leonie Gibson for technical assistance and Lynn Corcoran, Tom Gonda, Ora Bernard and Elizabeth Webb for discussions. This work was supported by the National Health and Medical Research Council, Canberra, the US NCI (CA12421), the American Heart Association and the Drakensberg Trust.

Received 4 December 1984; accepted 18 February 1985.

- Klein, G. *Nature* **294**, 313–318 (1981).
- Hayward, W., Neel, B. & Astrin, S. *Nature* **290**, 475–480 (1981).
- Payne, G., Bishop, J. M. & Varmus, H. *Nature* **295**, 209–214 (1982).
- Klein, G. *Cell* **32**, 311–315 (1983).
- Perry, R. *Cell* **33**, 647–649 (1983).
- Leder, P. *et al. Science* **222**, 765–771 (1983).
- Corcoran, L., Adams, J. M. & Cory, S. *Cell* **37**, 113–122 (1984).
- Li, Y., Holland, C., Hartley, J. & Hopkins, N. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6808–6811 (1984).
- Selten, G., Cuypers, H., Zijlstra, M., Melief, C. & Berns, A. *EMBO J.* **3**, 3215–3222 (1984).
- Steffen, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2097–2101 (1984).
- Tsichlis, P., Straus, P. & Hu, L. *Nature* **302**, 445–449 (1983).
- Lemay, G. & Jolicœur, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 38–42 (1984).
- Cuypers, H. T. *et al. Cell* **37**, 141–150 (1984).
- Tsichlis, P., Straus, P. & Kozak, C. *Molec. cell. Biol.* **4**, 997–1000 (1984).
- Tsichlis, P., Hu, L. & Strauss, P. G. in *UCLA Symposium: Normal and Neoplastic Hematopoiesis* (eds Golde, D. & Marks, P.) 399–415 (1983).
- Webb, E., Adams, J. M. & Cory, S. *Nature* **312**, 777–779 (1984).
- Cory, S., Graham, M., Webb, E., Corcoran, L. & Adams, J. *EMBO J.* (in the press).
- Ohno, S. *et al. Cell* **18**, 1001–1007 (1979).
- Ohno, S. *et al. J. exp. Med.* **159**, 1762–1777 (1984).
- Van Ness, B. *et al. Nature* **301**, 425–427 (1983).
- Cory, S., Gerondakis, S. & Adams, J. *EMBO J.* **2**, 697–703 (1983).
- Gerondakis, S., Cory, S. & Adams, J. M. *Cell* **36**, 973–982 (1984).
- Klein, E. & Klein, G. *J. natn. Cancer Inst.* **32**, 547–568 (1964).

- Famulari, N. *Curr. Topics Microbiol. Immun.* **103**, 75–108 (1983).
- Spira, J. *et al. Int. J. Cancer* **28**, 785–798 (1981).
- Cory, S. *et al. EMBO J.* **2**, 213–216 (1983).
- Chattopadhyay, S. *et al. Nature* **295**, 25–31 (1982).
- Herr, W. & Gilbert, W. *J. Virol.* **46**, 70–82 (1983).
- van der Putten, H. *et al. Cell* **24**, 729–739 (1981).
- Kelly, M., Holland, C., Lung, M. & Chattopadhyay, N. *J. Virol.* **45**, 291–298 (1983).
- Varmus, H. *Science* **216**, 812–820 (1982).
- Kozak, C. in *Mechanisms of B Cell Neoplasia* (eds Melchers, F., Potter, M. & Weigert, M.) 126–133 (Roche, Basle, 1983).
- Nusse, R., van Doyen, A., Cox, D., Fung, Y. & Varmus, H. *Nature* **307**, 131–136 (1984).
- Kozak, C., Sears, J. & Hogan, M. *Science* **221**, 867–869 (1983).
- Hollis, G. *et al. Nature* **307**, 752–755 (1984).
- Taub, R. *et al. Cell* **37**, 511–520 (1984).
- Davis, M., Malcolm, S. & Rabbitts, T. *Nature* **308**, 286–288 (1984).
- Croce, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 6922–6926 (1983).
- Erikson, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 7581–7585 (1983).
- Adams, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6966–6970 (1982).
- Adams, J., Gerondakis, S., Webb, E., Corcoran, L. & Cory, S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1982–1986 (1983).
- Mushinski, J. F., Bauer, S., Potter, M. & Reddy, E. P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1073–1077 (1983).
- Kelly, K., Cochran, B., Stiles, C. & Leder, P. *Cell* **25**, 603–610 (1984).
- Keath, E., Kelekar, A. & Cole, M. *Cell* **37**, 521–528 (1984).
- Gonda, T., Sheiness, D. & Bishop, J. M. *Molec. cell. Biol.* **2**, 617–624 (1982).
- Thomas, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

A novel serine esterase expressed by cytotoxic T lymphocytes

Mark S. Pasternack*† & Herman N. Eisen*‡

* Department of Biology and Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Children's and Medical Services (Infectious Disease Units), Massachusetts General Hospital, and the Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02114, USA

Cytotoxic T (T_c) lymphocytes recognize and lyse target cells and are thought to serve as an important defence against viral infections¹ and possibly against neoplasms². The nature of the receptors responsible for antigen recognition by these cells is becoming clearer, but the molecular mechanisms responsible for their cytolytic activity remain largely unknown. The possibility that proteases are involved in this process has been suggested by the effects of certain inhibitors^{3,4}. Here we demonstrate that clones of murine T_c cells possess considerable trypsin-like esterase activity when assayed by a sensitive colorimetric assay. This activity was blocked completely by two serine esterase inhibitors, diisopropylfluorophosphoridate (DFP) and phenylmethylsulphonyl fluoride (PMSF), but not by N^α -tosyl lysyl chloromethyl ketone (TLCK). The use of 3H -DFP as an affinity-labelling reagent demonstrated that the esterase activity resides in a protein of relative molecular mass (M_r) 28,000 (28K). A wide variety of other lymphocytes, including those from thymus, spleen and lymph node, established lines of B cells and noncytotoxic T cells, and clones of T helper cells, had about 300-fold less esterase activity than the T_c -cell clones and far smaller amounts of the DFP-reactive 28K protein. However, in thymocytes the esterase activity increased 20–50-fold and the 28K protein became more prominent 4 days after these cells had been stimulated *in vitro* to generate T_c cells.

Previous investigations have demonstrated that the lytic activity of both T_c cells^{3,4} and natural killer cells^{5–7} is reduced by exposure to certain protease inhibitors. In an attempt to

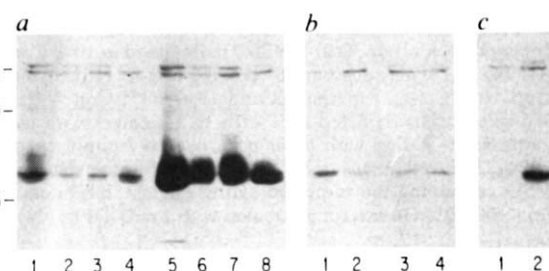


Fig. 1 Affinity labelling of serine esterase with 3H -DFP. **a**, Affinity labelling of T_c cells. Three million T_c cells were labelled directly (lanes 1–4) or following solubilization in NP40–PBS (lanes 5–8). Lanes 1, 5, G4; lanes 2, 6, B10; lanes 3, 7, 1.5.2; lanes 4, 8, cr 15 (2 days exposure). **b**, Affinity labelling of T_h -cell extracts. Three million cell equivalents were analysed. Lane 1, O_3 ; lane 2, D_5 ; lane 3, A_6A_2 ; lane 4, BW5147 (3 days exposure). **c**, Affinity labelling of thymocyte extracts. Six million cell equivalents were analysed. Lane 1, resting thymocytes; lane 2, thymocytes incubated for 4 days in Con A supernatant and Con A (3 days exposure). M_r standards ($\times 10^{-3}$) are shown on the left.

Methods. [3H]-DFP in propylene glycol (5.2 Ci mmol⁻¹; NEN) was added to 300 μ l of various cell extracts in NP40–PBS as described in Table 1 to a final concentration of 10^{-5} M. After 30 min incubation at 37 °C, the reaction was stopped by adding Tris-HCl pH 8.1, to a final concentration of 15 mM and SDS to a final concentration of 0.1%. Samples were centrifuged at 8,000g for 15 min. The pellets were discarded, the supernatants were concentrated to 100–150 μ l and the labelled proteins recovered by cold acetone precipitation. The protein pellet was dissolved in sample buffer containing 5% mercaptoethanol and analysed by electrophoresis on a 10% polyacrylamide gel in the presence of 0.1% SDS (ref. 19). Gels were impregnated with Autofluor (National Diagnostics, Somerville, New Jersey), dried and exposed for the indicated times at –70 °C using XAR-5 film (Eastman Kodak).

characterize T_c -cell proteases directly, extracts of G4 cells, a clone of BALB.B anti-H-2D^d T_c cells⁸, were prepared in phosphate-buffered saline (PBS) containing 0.5% Nonidet P-40 (NP40). Trypsin-like esterase activity in the lysate was assayed spectrophotometrically using a sensitive coupled reaction involving N^α -benzyloxycarbonyl-L-lysine thioester

† To whom correspondence should be addressed.

Table 1 Esterase activity in extracts of cytotoxic T lymphocytes and other cells

Cells	Units per 10 ⁶ cell equivalents	Cells	Units per 10 ⁶ cell equivalents
Cytotoxic T-lymphocyte clones		Other	
1. G4	172-311	14. P815	0.518
2. G4 (IL-2)	203.4	15. Thymocytes	0.040
3. B10	67.4	16. Con A-activated thymocytes	3.47
4. 1.5.2	84.3	17. Resting spleen cells	0.557
5. cr 15	77.6	18. Splenic B cells	0.323
Noncytotoxic T lymphocytes		19. Splenic LPS blasts	0.136
6. O ₃	0.520	20. B-cell clone 670.6	0.259
7. D ₅	0.431	21. B-cell clone 670.11	0.251
8. A ₆ A ₂	0.202	22. B-cell clone 750.17	0.227
9. EL-4	0.572	Inhibition studies	
10. EL-4 (RCASUP)	0.800	% Inhibition	
11. BW5147	0.258	23. G4 (DFP-treated)	98.7
12. BW5147 (RCASUP)	0.268	24. G4 (PMSF-treated)	97.2
13. BW5147 (RCASUP+Con A)	0.365	25. G4 (TLCK-treated)	4.6

Cell preparations containing non-viable stimulator cells or complement-lysed cells were passed over Ficol-Hypaque gradients (Pharmacia) to remove dead cells before extraction of the viable cells. G4 cells, which are adherent, were collected by brief exposure to EDTA (5 mM in PBS) after decanting stimulator cell debris; they were not purified further. Cells were washed twice in fresh medium, twice in PBS or in RPMI 1640 containing 100 µg ml⁻¹ bovine serum albumin (5×recrystallized; Boehringer-Mannheim), counted, washed a third time in PBS, then lysed by incubating 1–2×10⁷ cells ml⁻¹ in PBS containing 0.5% NP40 (Particle Data Laboratories, Elmhurst, Illinois) for 20–30 min on ice with frequent vortexing. Esterase activity in the lysates was measured by adding 900 µl of a reaction mixture (0.2 M Tris-HCl pH 8.1, 2×10⁻⁴ M BLT (Calbiochem-Behring), 2.2×10⁻⁴ M dithiobis (nitrobenzoic acid); (Sigma) to 100 µl of appropriate dilutions of cell extracts in NP40/PBS. After 30 min at room temperature, the absorbance at 412 nm was measured in a Gilford spectrophotometer using 100 µl 0.5% NP40 in PBS plus 900 µl of the reaction mixture as a blank; an absorbance of 1.0 was defined as 1 unit of esterase activity. The cell extracts were prepared from (1) G4, a BALB.B anti-H-2D^d T_c-cell clone (ref. 8); (2) G4 grown in recombinant human IL-2; (3) B10, a BALB.B anti-H-2L^d T_c-cell clone (ref. 8); (4) 1.5.2, a BALB.B anti-H-2L^d T_c-cell; (5) cr 15, a (B10.BR×B10.D2)F₁ anti-BALB minor T_c-cell clone (ref. 15); (6) O₃, a BALB/c T_H-cell clone specific for ovalbumin plus Ia^d (ref. 16); (7) D₅, a (BALB/c×A/J)F₁ T_H-cell clone specific for arsanilated protein plus Ia^d (ref. 17); (8) A₆A₂, a T_H-cell hybridoma of B10.A origin fused with BW5147, specific for hen egg lysozyme (provided by Dr L. Glimcher, manuscript in preparation); (9) EL-4, a T-cell leukaemia of C57BL/6 origin; (10) EL-4 maintained in 10% rat spleen cell Con A supernatant (RCASUP); (11) BW5147, T-cell lymphoma of AKR origin; (12) BW5147 maintained in 10% Con A supernatant; (13) BW5147 maintained in 10% Con A supernatant and 10 µg ml⁻¹ Con A; (14) P815, mastocytoma of DBA/2 origin; (15) thymocytes collected from 3-week-old BALB/c females; (16) thymocytes cultured for 4 days with 10% Con A supernatant and 10 µg ml⁻¹ Con A (ref. 10); (17) BALB/c spleen cells following lysis of erythrocytes with NH₄Cl; (18) BALB/c spleen cells depleted of T cells by treatment with two anti-Thy-1 monoclonal antibodies and rabbit complement; (19) BALB/c spleen cells cultured for 4 days with *Salmonella typhosa* lipopolysaccharide (LPS, Difco) at 10 µg ml⁻¹, then depleted of T cells with anti-Thy-1 before assay; (20–22) B-cell clones of CBA/N mice¹⁸. For the inhibition studies, an aliquot of G4 extract was incubated for 30 min at 37 °C with DMSO, or DMSO containing the indicated serine esterase inhibitors. Percentage inhibition = (1 – absorbance of treated samples/absorbance of control sample) × 100. (23) G4 extract pretreated with 2 mM DFP; (24) G4 extract pretreated with 2 mM PMSF; (25) G4 extract pretreated with 5×10⁻⁴ M TLCK.

(BLT) as described by Green and Shaw⁹; hydrolysis yields benzyl mercaptan, which reacts with Ellman's reagent to produce the thiophenoxide chromophore (Table 1). The esterase activity, equivalent to ~1 µg of trypsin per 10⁶ cells, was not restricted to G4, as three other T_c-cell clones (B10, 1.5.2 and cr 15), of two different specificities, had roughly comparable activity (Table 1).

Because all of the T_c-cell clones had been maintained in long-term culture in medium containing 10% rat spleen cell concanavalin A (Con A) supernatant and stimulator cells, it was important to assay the stimulator cells and a variety of other cell lines. Table 1 summarizes the BLT-esterase activities in NP40 extracts of several different noncytotoxic cell lines. The cells used as stimulators to maintain the cultured T_c-cell clones (BALB/c spleen cells and *in vitro* passaged P815) possessed little activity. Two T-cell tumour lines, EL-4 and BW5147, also had minimal activity. Furthermore, when these T-cell lines were maintained in medium containing 10% Con A supernatant (or when BW5147 was incubated with both 10% Con A supernatant and 10 µg ml⁻¹ Con A) there was no significant increase in enzyme activity. Thus, the presence of stimulator cells and conditioned medium as sources of adsorbed esterase activity could not be invoked as trivial explanations.

Two additional experiments confirmed the endogenous character of the esterase activity in T_c cells. First, there was no difference between the esterase activity in G4 cells that had been serially cultured twice in recombinant human interleukin-2 (IL-2, a generous gift of Biogen) and in cells cultured continuously in 10% Con A supernatant. Second, and more interestingly, two T-helper (T_H) cell clones, O₃ and D₅, which lacked cytotoxic

activity and had been maintained in long-term culture with 10% Con A supernatant and intermittently stimulated with antigen-pulsed spleen cells, had negligible activity (Table 1). A T_H-cell hybridoma, A₆A₂, also showed only traces of activity (Table 1). Thus, among all B and T lymphocytes tested, a high level of esterase activity was expressed only in T_c cells.

To determine whether the esterase activity in T_c-cell clones represented some unusual adaptation of these cells to long-term culture, we studied the expression of esterase activity in T_c cells generated by short-term culture of precursor cells. For this purpose, thymocytes were incubated in culture medium containing rat spleen cell Con A supernatant and Con A¹⁰. After 4 days, viable thymocytes showed a level of lectin-dependent cytotoxic activity (2–4 lytic units (LU₅₀, see Fig. 2 legend) per 10⁶ cells)¹¹ that was comparable to that seen in conventional mixed lymphocyte cultures, and extracts of these thymocytes showed an approximately 50-fold increase in BLT-esterase activity (see Table 1).

The T_c-cell esterase activity was completely inhibited by the irreversible serine esterase inhibitors DFP and PMSF (Table 1), suggesting that this activity is due to a serine esterase, but it was not inhibited by TLCK, a trypsin inhibitor known to block T_c-cell cytotoxicity and to react with the surface glycoproteins T200 and LFA-1 (refs 3, 4, 8). Using ³H-DFP as an affinity-labelling reagent¹² to characterize the esterase further (Fig. 1), we found that although minor ³H-labelled components occurred at M_r 76,000 and 71,000, a prominent ³H-DFP-reactive component was present at M_r 28,000. Moreover, whereas the higher-M_r moieties were present at about the same low level in all lymphoid cells tested, the level of the 28K component clearly

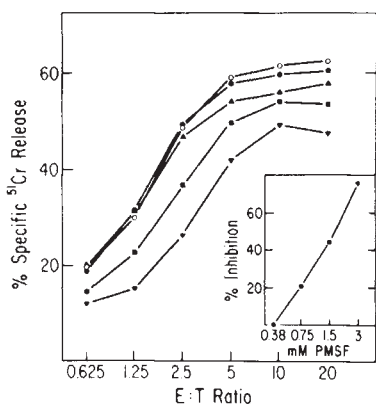


Fig. 2 Inhibition by PMSF of specific target cell lysis by T_c -cell clone G4. E:T, effector to target ratio. The G4 cells were suspended in complete medium at $\sim 5 \times 10^6 \text{ ml}^{-1}$ and DMSO (control) or PMSF dissolved in DMSO was added (final DMSO concentration 3%). ○, Control with DMSO alone; ●, 0.375 mM PMSF; ▲, 0.75 mM PMSF; ■, 1.5 mM PMSF; ▼, 3 mM PMSF. Inset: Results expressed as a dose-response curve. One lytic unit (LU_{50}) is the number of T_c cells required for 50% specific lysis of 1×10^4 ^{51}Cr -labelled P815 cells. Percentage inhibition = $(1 - LU_{50} \text{ in the presence of PMSF} / LU_{50} \text{ in the absence of PMSF}) \times 100$.

Methods. The cells were incubated with PMSF (or DMSO as control) for 30 min at 37°C , washed, counted and diluted for titration of cytotoxic activity against 1×10^4 ^{51}Cr -labelled P815 cells. Recovery of G4 was comparable for all the treated samples. In some experiments, DMSO alone was responsible for ~ 10 – 15% ^{51}Cr release from labelled G4 cells, but the presence of PMSF at the concentrations shown had no additive toxicity measured either by ^{51}Cr release or by reduced recovery of viable (trypan blue-negative) G4 cells. T_c -cell assays were performed as described previously²⁰.

varied with the level of the BLT-esterase activity; both were pronounced in extracts of T_c -cell clones (Fig. 1a, lanes 5–8), but present in only trace amounts in noncytotoxic cells (Fig. 1b, lanes 1–4), and both increased after inducing T_c -cell activity in thymocytes by incubation with Con A (Fig. 1c, lanes 1, 2). Further evidence linking the BLT-esterase activity to the 28K component has emerged from their co-purification through affinity chromatography and an ion-exchange column (M. S. Pasternack *et al.*, manuscript in preparation).

If BLT-esterase is necessary for target cell lysis, inhibition of the esterase should block the cytolytic reaction. DFP was not tested because propylene glycol, the solvent in which it is kept, is toxic to cells, but PMSF at nontoxic concentrations clearly inhibited lysis of P815 ($H-2^d$) cells by the cloned anti- D^d G4 cells (Fig. 2). Note, however, that although PMSF at 2 mM blocked the esterase activity in lysates completely, at 1.5 mM it blocked only 50% of the T_c -cell cytolytic activity (Table 1, Fig. 2). Similarly, ^3H -DFP labelled the 28K component intensely in T_c -cell lysates, but only weakly in intact T_c cells (Fig. 1a, compare lanes 1–4 with 5–8). The difference between lysed and intact T_c cells raises several possibilities: one is that the BLT-esterase, or at least its active site, is normally located intracellularly, rather than on the cell surface, and is thus only poorly accessible to the added esterase inhibitors; another is that the esterase exists in intact cells as an inactive zymogen until converted into an active enzyme by another protein, perhaps a member of a complement-like cascade. In this connection, it is of interest that structures resembling the aggregated C9 component of complement have recently been described in cytolytic cells (T_c cells and natural killer cells)^{13,14}. We believe the BLT-esterase described here represents the first enzymatic difference reported between cytotoxic and noncytotoxic T cells.

We thank Michael Bevan for cr 15 cells, Jonathan Braun for B-cell clones, Rosemarie DeKruyff for T_h cells O₃ and D₅, Laurie Glimcher for hybridoma A₆A₂, and David Kranz and David Sherman for clone 1.5.2. We also thank Ann Hicks for assistance in the preparation of this manuscript. This work was supported

by research grants (CA-15472 and CA-28900) and a Cancer Center Core Grant (CA-14051) from the National Cancer Institute.

Note added in proof: Preliminary results suggest that cloned T cells having both the helper phenotype ($L3T4^+$, $Lyt-2^-$, antigen-dependent IL-production) and exhibiting Con A-dependent cytotoxic activity may also have elevated serine esterase activity. Thus it is possible that the serine esterase level correlates with cytotoxic activity *per se*, rather than with the surface glycoprotein markers that classically distinguish helper (T_H) and cytotoxic (T_C) cells.

Received 24 December 1984; accepted 11 February 1985.

1. Lukacher, A. E., Braciale, V. L. & Braciale, T. J. *J. exp. Med.* **160**, 814–824 (1984).
2. Engers, H. D., Glasebrook, A. L. & Sorenson, G. D. *J. exp. Med.* **156**, 1280–1285 (1982).
3. Chang, T.-W. & Eisen, H. N. *J. Immun.* **124**, 1028–1033 (1980).
4. Redelman, D. & Hudig, D. *J. Immun.* **124**, 870–878 (1980).
5. Hudig, D., Haverty, T., Fulcher, C., Redelman, D. & Mendelsohn, J. *J. Immun.* **126**, 1569–1574 (1981).
6. Quan, P.-C., Ishizaka, T. & Bloom, B. R. *J. Immun.* **128**, 1786–1791 (1982).
7. Hudig, D., Redelman, D. & Minning, L. *J. Immun.* **133**, 2647–2654 (1984).
8. Pasternack, M. S., Sitkovsky, M. V. & Eisen, H. N. *J. Immun.* **131**, 2477–2483 (1983).
9. Green, G. & Shaw, E. *Analyt. Biochem.* **93**, 223–226 (1975).
10. Irle, C., Piquet, P.-F. & Vassalli, P. *J. exp. Med.* **148**, 32–45 (1978).
11. Cerottini, J. C. & Brunner, K. *Adv. Immun.* **18**, 67–132 (1974).
12. Heck, L. W., Remold-O'Donnell, E. & Remold, H. G. *Biochem. biophys. Res. Commun.* **83**, 1576–1583 (1978).
13. Podack, E. R. & Konigsberg, P. J. *J. exp. Med.* **160**, 695–710 (1984).
14. Henkart, P. A., Millard, P. J., Reynolds, C. W. & Henkart, M. P. *J. exp. Med.* **160**, 75–93 (1984).
15. Hunig, T. R. & Bevan, M. J. *J. exp. Med.* **155**, 111–125 (1982).
16. Clayberger, C., DeKruyff, R. H., Fay, R., Pavlakis, M. & Cantor, H. *J. Immun.* **133**, 1174–1178 (1984).
17. Rao, A., Faas, S. J. & Cantor, H. *J. exp. Med.* **159**, 479–494 (1984).
18. Braun, J. *J. Immun.* **130**, 2113–2116 (1983).
19. Laemmli, U. K. *Nature new Biol.* **227**, 680–684 (1975).
20. Celis, E., Hale, A. H., Russell, J. H. & Eisen, H. N. *J. Immun.* **122**, 954–958 (1979).

Reversible activation of non-steroid binding oestrogen receptor

William J. Raymoure*, Reid W. McNaught* & Roy G. Smith*†

Roy and Lillie Cullen Department of Urological Research, Departments of *Urology and †Cell Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA

Two high-affinity oestrogen receptors have been identified in the chick oviduct with equilibrium dissociation constants (K_d) of 0.1 and 1 nM, differing in their binding kinetics, role in ovalbumin synthesis and independent regulation *in vivo*. The higher-affinity receptor (X) increases RNA polymerase II activity directly¹, whereas the low-affinity receptor (Y) seems to be necessary to confer specificity to transcription of oestrogen-dependent genes². Acute administration of progesterone to oestrogen-stimulated chicks results in preferential destruction of the nuclear Y receptor³ accompanied by interruption of ovalbumin gene transcription⁴. Here we demonstrate that receptor Y exists in a non-oestradiol binding form (Y_{nb}) which can be activated to the binding form *in vitro* by treatment with either ATP or ADP. Furthermore, dialysis of oviduct cytosol, which has no effect on the high-affinity receptor X, converts receptor Y to Y_{nb} ; receptor Y can then be recovered by treatment with ATP in the presence of Mg^{2+} and independently of Ca^{2+} . This is the first report of the controlled interconversion between a non-steroid binding form of oestrogen receptor and active receptor in a tissue that contains two independently regulated oestrogen receptor types.

Oestrogen receptor augmentation by incubation with nucleotides has been reported by Fleming *et al.*⁵ and Auricchio *et al.*⁶. However, the affinity and steroid specificity of the induced binding sites were not examined. The characteristics of the heterogeneity of oestrogen receptors described in the chick oviduct^{7–9} are different from the oestrogen-binding proteins of the rat uterus¹⁰. To avoid comparisons with the rat uterine types I and II oestrogen-binding proteins, the two oviduct receptors have been designated X ($K_d = 0.1$ nM) and Y ($K_d = 1$ nM).

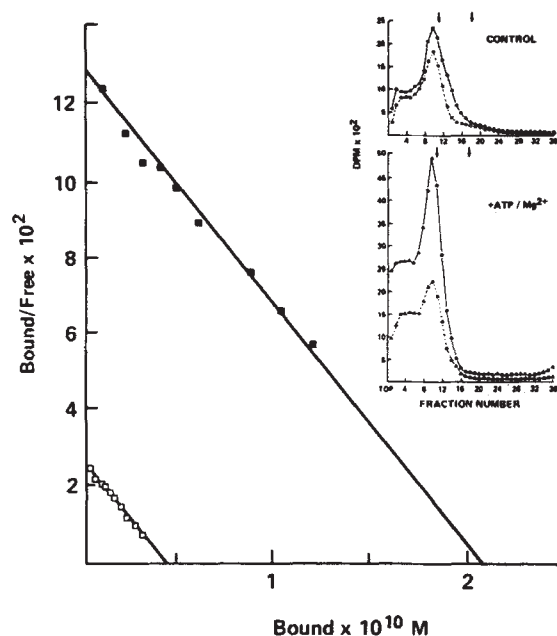


Fig. 1 Effects of ATP on the Y-receptor type. Week-old white Leghorn chicks were stimulated for 14 days with diethylstilboestrol in sesame oil and withdrawn for 1–2 days. Oviducts were removed into ice-cold saline and homogenized in TESH buffer (10 mM Tris, 1 mM EDTA, 12 mM thioglycerol, pH 7.4) in a ratio of 1/5 (w/v) and centrifuged at 1,000g for 20 min. Cytosol was obtained from the supernatant by centrifugation at 100,000g for 90 min. Binding assays used a range of $^3\text{H-E}_2$ of 80 pM–10 nM and included 10 mM Na_2MoO_4 to stabilize the binding form of the receptor. Nonspecific binding was assessed by addition of 100-fold excess diethylstilboestrol to duplicate tubes. Assays were incubated for 3 h at 30 °C, then for 18 h at 4 °C; unbound steroid was separated by addition of dextran-coated charcoal⁴ and data analysed according to Scatchard¹⁴. $\square$ — $\square$, Control assay ($K_d = 1.5$ nM; $n = 0.43 \times 10^{-10}$ M); $\blacksquare$ — $\blacksquare$, inclusion of 2.5 mM ATP and 2.5 mM MgCl_2 in the assay ($K_d = 1.6$ nM; $n = 2.0 \times 10^{-10}$ M). Insets: Sucrose density gradient analysis. Aliquots of cytosol used in binding assays were labelled with 5 nM ^3H -oestradiol alone or with the addition of 2.5 mM ATP and 2.5 mM MgCl_2 at 30 °C for 3 h then at 4 °C for 18 h to augment receptor Y. Nonspecific binding was assessed by the inclusion of 100-fold molar excess unlabelled diethylstilboestrol. Unbound steroid was removed by the addition of dextran-coated charcoal immediately before the cytosol was applied to linear gradients containing 5–20% sucrose (w/v) with 400 mM KCl. Gradients were centrifuged for 9 h at 200,000g with ^{14}C -ovalbumin (3.6S) and ^{14}C -amylase (8.2S) included as internal standards. Upper inset: Control, total bound ($\bullet$ — $\bullet$); nonspecific ($\circ$ — $\circ$). Lower inset: Augmented Y, total bound ($\blacktriangle$ — $\blacktriangle$); nonspecific ($\triangle$ — $\triangle$).

Oviduct cytosol, which contained low concentrations of receptor Y only, was obtained from chicks chronically stimulated with diethylstilboestrol for 14 days and then withdrawn for 1–2 days. Treatment of this cytosol with 2.5 mM ATP and 2.5 mM MgCl_2 in the presence of oestradiol at 30 °C for 3 h caused a large increase in the receptor concentration, as illustrated by Scatchard and sucrose density analyses (Fig. 1). This amplification was of a comparable degree when ADP was substituted for ATP, whereas GTP had a lesser effect. Both ATP with an ATP regenerating system (creatine kinase and creatine phosphate) and ADP with an ADP trap (creatine kinase and creatine) were effective. Incubation with the blocked ATP and ADP derivatives β, γ -methyleneadenosine 5'-triphosphate (AMP-PCP) and α, β -methyleneadenosine 5'-diphosphate (AMP-CP) did not increase the concentrations of Y, whereas α, β -methyleneadenosine 5'-triphosphate (AMP-CPP) was as potent as either ATP or ADP. Neither cyclic AMP nor cyclic GMP in the presence of isobutylmethylxanthine converted Y_{nb} to Y, nor did they alter the ATP effect. The conversion of Y_{nb} to Y by ATP was dependent specifically on Mg^{2+} , with an optimum concentration of 2–5 mM. Mn^{2+} was ineffective and, in contrast to the observations of Auricchio *et al.*⁶, Ca^{2+} was ineffective and did not synergize the

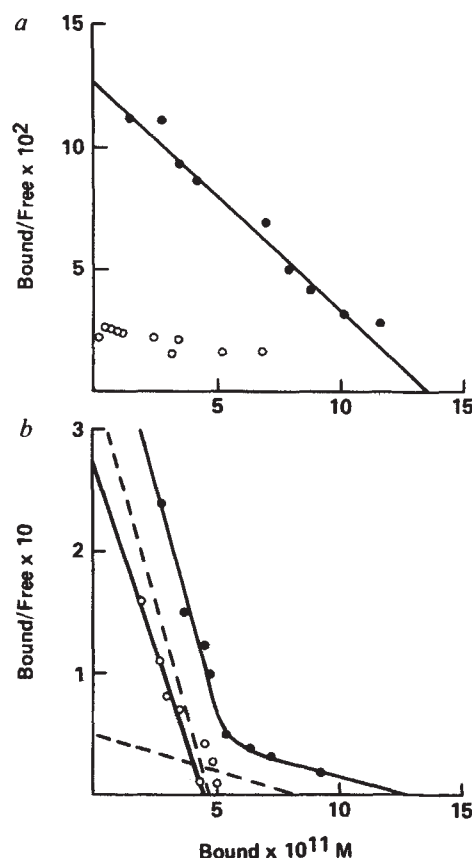


Fig. 2 Preferential effects of dialysis on the Y-receptor type. Chick oviduct cytosol was prepared, assayed and analysed as described in Fig. 1. **a**, Oviduct cytosol from chicks withdrawn from diethylstilboestrol for 3 days containing the Y-receptor type only. Closed circles indicate control assay ($K_d = 1.1 \times 10^{-9}$ M, $n = 13.3 \times 10^{-11}$ M); open circles denote dialysis of cytosol for 24 h. **b**, Oviduct cytosol from chicks withdrawn from diethylstilboestrol for 5 days containing both X- and Y-receptor types. Closed circles, control assay; dashed line, Rosenthal projection¹⁵ of data; $K_d X = 0.13 \times 10^{-9}$ M, $n = 4.3 \times 10^{-11}$; $K_d Y = 1.5 \times 10^{-9}$ M, $n = 7.9 \times 10^{-11}$ M; open circles, assay data from cytosol after dialysis for 24 h as in Fig. 1. $K_d X = 0.13 \times 10^{-9}$ M, $n = 4.3 \times 10^{-11}$ M.

effect of Mg^{2+} . The augmentation did not occur in cytosol obtained from oviducts withdrawn from diethylstilboestrol for >3 days when a high concentration of X, as well as Y, was present. The data shown in Table 1 demonstrate that the augmented receptor has steroid specificity typical of oestrogen receptors.

The Y-receptor concentration can be reduced selectively by dialysis (Fig. 2). Oviduct cytosol containing only the Y-receptor type was obtained from chicks withdrawn from diethylstilboestrol after 2 days. After this cytosol was dialysed for 24 h against TESH buffer (see Fig. 1 legend), oestrogen receptor was no longer detectable (Fig. 2a). Cytosol from chicks withdrawn from diethylstilboestrol for longer times (3–5 days) with both the X and Y receptors was also dialysed; no effect was observed on receptor X, whereas Y was decreased preferentially to undetectable levels (Fig. 2b).

Evidence for the reversible *in vitro* conversion of Y to Y_{nb} is presented in Table 2. Dialysis of cytosol caused a complete loss of the Y-receptor type. When dialysed cytosol was treated with ATP and MgCl_2 , there was essentially quantitative recovery ($97.8 \pm 1.1\%$, $n = 3$) of the Y receptor. Treatment of dialysed cytosol with either MgCl_2 or ATP alone did not restore lost binding activity.

This study provides the first evidence that the chick oviduct oestrogen receptor Y exists in a non-oestradiol binding form, Y_{nb} , and that these two forms can interconvert; dialysis of cytosol favours Y_{nb} ; Y_{nb} can be converted to Y at 30 °C in the

Table 1 Specificity of oestradiol binding to receptor Y after augmentation with ATP

Steroid (100 nM)	³ H-Oestradiol bound (pM) (% inhibition)
None	235.59 (0)
Oestradiol	124.77 (47.0)
Oestriol	163.43 (30.6)
Diethylstilboestrol	130.42 (44.7)
Progesterone	229.14 (2.7)
Testosterone	228.74 (2.9)
Dihydrotestosterone	227.79 (3.3)
Cortisol	236.06 (0)
Triamcinolone acetoneide	237.82 (0)

Single-point (1 nM ³H-oestradiol (E₂)) binding assays were performed in the presence of various unlabelled steroids in conditions of receptor augmentation as described in Fig. 1 legend.

presence of MgCl₂ and ATP, ADP or GTP. The Mg²⁺-dependence of this conversion, the requirement of a hydrolysable terminal phosphate group for blocked nucleotide analogues to elicit conversion and the ineffectiveness of cyclic nucleotides (even with added ATP) suggest that a non-cyclic-nucleotide-mediated phosphorylation is involved. That both ATP and ADP, together with their regenerating and trap systems, respectively, augment receptor activity suggests that either ATP or ADP can donate the necessary phosphate group. The reversibility of this conversion (Table 2), the lack of receptor augmentation with the hydrolysis-resistant ATP and ADP derivatives and the effectiveness of ATP with regenerating system, eliminates the possibility that these observations are an ADP-mediated protection against heat loss, as has been suggested for glucocorticoid receptors¹¹.

We have been unable to observe the non-oestrogen binding form of the receptor in oviduct cytosol preparations from either chronically stimulated or long-term (> 3 days) oestrogen-withdrawn chicks. Because the highest concentration of Y_{nb} occurred when the cytoplasmic receptor concentration was lowest, it seems plausible that the non-binding form represents a recycling pool of receptor. It is also conceivable that Y_{nb} represents the first stage of irreversible receptor degradation. We do not favour the latter explanation for the following reasons: (1) the conversion of the non-binding to the oestrogen-binding form is possible simply by increasing the concentration of Mg²⁺ and specific nucleotides in the cytosol; (2) sucrose gradient analysis indicates that the receptor obtained by conversion from Y_{nb} has similar physical properties to the endogenous steroid-binding form; (3) although soluble receptor concentrations vary according to the time after withdrawal from diethylstilboestrol, the concentration of receptors after treatment with ATP/Mg²⁺ shows much less variation—after withdrawal from diethylstilboestrol, the concentration of Y_{nb} is inversely proportional to that of the oestradiol-binding receptors. Collectively, these data indicate that the controlled reversible interconversion between Y and Y_{nb} is

Table 2 Reversibility of the conversion of oestrogen receptor Y to its non-oestradiol binding form

Treatment	Concentration of Y × 10 ¹⁰ M	K _d (nM)
Control	0.81	0.91
Dialysis	Undetectable	0
Dialysis, ATP/Mg ²⁺	0.78	1.0

Binding assays were performed on chick oviduct cytosol and analysed as described in Fig. 1. Cytosol was dialysed against two changes of 1,000-fold excess TESH buffer for 24 h at 4 °C using Spectrapor dialysis tubing with a relative molecular mass cutoff of 3,500. ATP and Mg²⁺ were used in the presence of an ATP regenerating system according to the method of Barnett *et al.*¹¹. Assays were incubated at 30 °C for 3 h then at 4 °C for 18 h. In three experiments, cytosol dialysis reduced Y concentration to undetectable levels; addition of ATP/Mg²⁺ resulted in recovery of 97.8 ± 1.1% of Y.

associated with physiological receptor recycling. Furthermore, because recovery of Y receptors (which have been reduced *in vivo* or by dialysis treatment *in vitro*) can be accomplished by treatment of Y_{nb} with Mg²⁺ and ATP, the continued presence of ATP and Mg²⁺ may be necessary for maintenance of the steroid-binding characteristics of the Y form of the oestrogen receptor, similar to the mechanism proposed for the glucocorticoid receptor by Munck *et al.*¹². The precise mechanism of these intriguing observations is currently under investigation.

Oestrogen receptor augmentation in the calf uterus, which contains only one form of active oestrogen receptor, has been shown recently by Migliaccio *et al.*¹³ to be associated with a Ca²⁺-stimulated phosphorylation. Although we agree that a phosphorylation reaction is associated with receptor augmentation, we do not find that Ca²⁺ stimulates this process in our system.

We thank Nita Burke for technical assistance. This work was supported by NIH grant HD17727 and in part by Robert A. Welch Foundation grant AU844.

Received 2 November 1984; accepted 5 February 1985.

1. Taylor, R. N. & Smith, R. G. *Biochemistry* **21**, 1791-1787 (1982).
2. Smith, R. G. & Taylor, R. N. *J. Steroid Biochem.* **15**, 321-328 (1981).
3. McNaught, R. W., Raymoure, W. J. & Smith, R. G. *Endocrinology* **112**, Suppl., 342 (1983).
4. Seaver, S. S., Hoffman, J. F. & Coulson, P. B. *J. Steroid Biochem.* **13**, 1269-1276 (1980).
5. Fleming, H., Blumenthal, R. & Gurpide, E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2486-2490 (1983).
6. Auricchio, F., Migliaccio, A., Castoria, G., Latoria, S. & Schiavone, E. *Biochem. biophys. Res. Commun.* **101**, 1171-1178 (1981).
7. Smith, R. G., Clarke, S. G., Zalta, E. & Taylor, R. N. *J. Steroid Biochem.* **10**, 31-35 (1979).
8. Taylor, R. N. & Smith, R. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1742-1746 (1982).
9. Taylor, R. N., Swaneck, G. E. & Smith, R. G. *Biochem. J.* **192**, 385-393 (1980).
10. Clark, J. H., Hardin, J. W., Upchurch, S. & Eriksson, H. *J. Biol. Chem.* **253**, 7630-7634 (1978).
11. Barnett, C. A., Palmour, R. M., Litwack, G. & Seegmiller, J. E. *Endocrinology* **112**, 2059-2068 (1983).
12. Munck, A. *et al. J. Steroid Biochem.* **3**, 567-578 (1972).
13. Migliaccio, A., Rotondi, A. & Auricchio, F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5921-5925 (1984).
14. Scatchard, G. *Ann. N.Y. Acad. Sci.* **51**, 660-672 (1949).
15. Rosenthal, H. E. *Analyt. Biochem.* **20**, 525-532 (1967).

Dihydropyridine receptors in muscle are voltage-dependent but most are not functional calcium channels

Lawrence M. Schwartz*, Edwin W. McCleskey* & Wolfhard Almers

Department of Physiology and Biophysics, University of Washington, Seattle, Washington 98195, USA

1,4-Dihydropyridines are a new class of compounds believed to bind specifically and with high affinity to voltage-dependent calcium channels¹⁻⁵. They may be the first example of a ligand of use in the extraction^{1,2} and purification³ of the Ca channel. Although Ca channels and dihydropyridine receptors are found in many tissues, the richest and most convenient source is skeletal muscle⁴. Functionally, 1,4-dihydropyridines such as nifedipine and nitrendipine block Ca channels⁶⁻¹¹; this effect is believed to form the basis for their clinical importance as Ca antagonists in relaxing vascular smooth muscle⁶. But where currents through Ca channels can be measured directly⁷⁻¹⁰, the block has required 100-1,000 times higher concentrations of dihydropyridine than necessary for the saturation of dihydropyridine binding sites^{1,5}. This discrepancy has remained unresolved because the study of pharmacological effects on Ca channels has required intact cells, while it has been difficult to investigate binding in other than cell-free preparations. Here we describe a method for measuring dihydropyridine binding to intact skeletal muscle and we compare our results with voltage-clamp measurements of Ca-channel block. We conclude that less than a few per cent of the binding sites in skeletal muscle represent functional Ca channels, contrary to general belief.

* Present addresses: Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27514, USA (L.M.S.); Department of Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA (E.W.McC.).

Figure 1a shows uptake of the 1,4-dihydropyridine ^3H -PN-200/110, a new and potent Ca antagonist^{12,13} and close analogue of nitrendipine, by intact frog sartorius muscle. While the drug was taken up, its concentration in the medium fell. The drug content of the muscles, calculated from the disappearance of drug from the medium, stabilized within 1–2 h. Unfortunately, most of the uptake is an unsaturable distribution of the highly lipophilic drug into the hydrophobic reservoirs of the muscle. To test for the presence of any saturable component of uptake, five pairs of frog sartorius muscles were incubated for 6 h in Ringer's solution containing ^3H -PN-200/110 at a final concentration of 0.6 nM. With each pair, the first muscle was exposed to labelled drug alone and the second exposed also to a saturating concentration (700 nM) of unlabelled nitrendipine. After the muscles were blotted, weighed, digested and assayed for radioactivity, it appeared that nitrendipine reduced the uptake of the label ^3H -PN-200/110 by an average of $8.0 \pm 0.8 \text{ pmol g}^{-1}$, indicating that the label and nitrendipine compete for a finite number of sites. However, a large and presumably nonsaturable uptake remained in nitrendipine (13.8 pmol g^{-1}) and made similar measurements at higher drug concentrations prohibitively noisy.

To diminish noise, a protocol was developed to measure saturable and nonsaturable components on the same muscle. The experiment in Fig. 1b is similar to that shown in a, but after equilibrium had been reached, a small volume of concentrated nitrendipine in ethanol was added to displace any saturably bound label (arrow). In the following 1–2 h, the muscle released label (1.9 pmol g^{-1}) and the concentration of ^3H -PN-200/110 in the medium increased from 1.6 to 2.1 nM (not shown). When ethanol is added without nitrendipine, release of the label is much less or does not occur.

The experiment may be used to determine the amount of saturably bound label. Before nitrendipine was added, the total uptake ($B_t = 63.9 \text{ pmol g}^{-1}$ in Fig. 1b) had saturable (B_s) and nonsaturable components ($B_{n,b}$), both in equilibrium with 1.6 nM label:

$$B_t = B_s + B_{n,b} \quad (1)$$

Two hours after adding nitrendipine, all label remaining in the muscle, $B_{n,a}$ (62.0 pmol g^{-1} in Fig. 1b), should be contained in the nonsaturable pool. Subscripts 'b' and 'a' refer to values before and after addition of nitrendipine. $B_{n,a}$ should be larger than $B_{n,b}$ because the muscle is now in equilibrium with a higher concentration of label. If nonsaturable binding increases linearly with concentration, then

$$B_{n,b} = B_{n,a}[\text{label}]_b/[\text{label}]_a \quad (2)$$

In Fig. 1b, $B_{n,b} = 62.0 \times 1.6/2.1 = 47.2 \text{ pmol g}^{-1}$; therefore, from equation (1), $B_s = 63.9 - 47.2 = 16.7 \text{ pmol g}^{-1}$. Thus, at 1.6 nM, the muscles in Fig. 1b bound 16.7 pmol g^{-1} of label that could be displaced by nitrendipine.

Figure 2A (curve a) was obtained from experiments similar to Fig. 1b and plots the amount of displaced label against the concentration measured just before the addition of nitrendipine. The curve describes one-to-one binding of label to a single class of sites. The parameters providing the best least-squares fit to the data are a dissociation constant (K_d) of $0.93 \pm 0.07 \text{ nM}$ and a binding capacity at infinite concentration (B_{max}) of 26 pmol g^{-1} (dashed line). Uptake was independent of whether ^3H -PN-200/110 was displaced by nitrendipine or by unlabelled PN-200/110 (five muscles). The filled circle was not included in the fit; it was obtained by counting muscles incubated with and without an excess of nitrendipine, as described above. Within experimental error, the result agrees with the other data on curve a.

When muscles were incubated in a potassium-rich solution (membrane potential -15 mV instead of the -88 mV observed in Ringer's), K_D for ^3H -PN-200/110 remained unchanged ($1.00 \pm 0.12 \text{ nM}$), but B_{max} was three times higher (79 pmol g^{-1}). The apparent increase in the number of binding sites cannot result from an increase in intracellular Ca^{2+} concentration

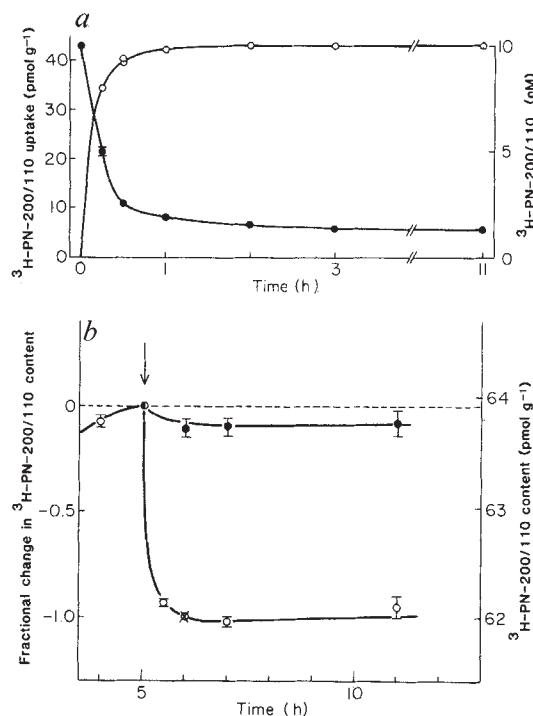


Fig. 1 a, Uptake of ^3H -PN-200/110 by intact frog sartorius muscles; b, release of ^3H -PN-200/110 by unlabelled nitrendipine. **Methods.** a, Each muscle (weight 40–110 mg) was incubated in a microcentrifuge tube containing 0.35 ml Ringer's solution (115 mM NaCl, 2.5 mM KCl, 1.8 mM CaCl_2 , 3 mM phosphate buffer, pH 7.2), ^3H -PN-200/110 at an initial concentration of 10 nM and a 5-mm long magnetic stirrer for gentle agitation. At various times, 0.075-ml aliquots were withdrawn to measure the concentration (●) of label remaining in solution. The drug content of each muscle was calculated from the disappearance of drug from the incubate, allowing for the fact that the volume was diminished by each sample. Drug content was normalized to that measured after a 1-h incubation; each point plotted (○) is the mean from five muscles. In this experiment, the average uptake after 1 h was $43 \pm 2.2 \text{ pmol per g wet weight}$; this defines the left ordinate. When no muscle was present, the concentration of label did not decline with time, indicating that no label was taken up by the microcentrifuge tubes. b, Protocol as in a, except that the first and second samples were always taken after 4 and 5 h of incubation, respectively. After the second sample (arrow), 1 μl ethanol containing either no drug (●) or 1.5 mM nitrendipine (gift of Dr A. Scriabine, Miles Institute; ○), was added to the incubation tube and additional samples taken at later times. The added nitrendipine is expected to result, at equilibrium, in a concentration of $\sim 700 \text{ nM}$, as the drug dilutes into an average nitrendipine space of 1.8 ml per muscle ($30.5 \pm 0.9 \text{ ml g}^{-1}$, $n = 10$, mean muscle weight 60 mg). This concentration is sufficient to displace all bound label, as the same amounts of label were released within 1 h of adding either 2 times smaller or 10-fold larger amounts of nitrendipine (not shown). The left-hand ordinate plots the change in muscle content of ^3H -PN-200/110 that occurred after the addition of ethanol and/or nitrendipine. Throughout, the change in content was normalized to that at 6 h (○, ⊗). The average change was $1.9 \pm 0.1 \text{ pmol g}^{-1}$ ($n = 20$); this value, together with the average content at 5 h (63.9 pmol g^{-1} in this group of muscles), defines the right-hand ordinate. ●, Not normalized. Temperature, $22\text{--}24^\circ\text{C}$ throughout; ^3H -PN-200/110 (NEN, specific activity 85 Ci mmol^{-1}) was the optically pure (+) stereoisomer. Vertical error bars indicate s.e.m. and are omitted if smaller than the size of the symbol.

caused by release from the sarcoplasmic reticulum or, possibly, by failure of the Na/Ca exchange in the absence of external Na. First, incubation began long after the K-contracture accompanying the depolarization had subsided. Second, the absence of Ca^{2+} inside and out had no effect on uptake (Table 1, experiment 1), confirming previous results on cell-free preparations¹. Apparently, depolarization acts like the drug diltiazem,

which, in cell-free¹³ or even solubilized skeletal muscle preparations², also increases the number of nitrendipine or PN-200/110 binding sites without changing their affinity. The diltiazem result indicates that dihydropyridine receptors may exist in a form which does not bind antagonist dihydropyridines, or in which their affinity for them is too low to be detected by the usual assays. Diltiazem is thought to convert low- into high-affinity receptors; we suggest that depolarization has the same effect.

Figure 2B shows that the nonsaturable uptake depends linearly on concentration, as expected. Results in Ringer's and in high-K solution are identical. The regression line drawn through the data points corresponds to a PN-200/110 space of 27.5 ml per g wet weight.

The value for K_D obtained here agrees well with results from cell-free skeletal muscle preparations (for PN-200/110, $K_D = 1.4$ nM in guinea pigs¹³; for nitrendipine, $K_D = 1.7$ nM in rabbits¹), as does the value for B_{max} . Because intact frog muscles have a protein content of 153 ± 10 mg per g wet weight (Lowry method, five muscles), B_{max} in depolarized muscle is 0.5 pmol per mg protein; in homogenized (and therefore depolarized) rabbit muscle, $B_{max} = 0.8$ pmol g⁻¹ (ref. 1). This agreement supports the validity of our assay for dihydropyridine binding in intact muscle.

Table 1 shows that adrenaline (0.01 mM) causes no significant change in PN-200/110 binding (expt 2), even though the hormone increases current through Ca channels in cardiac muscle¹⁴. Muscles bound less ³H-nitrendipine than ³H-PN-200/110 (ref. 13), but the preference for PN-200/110 is apparently less marked in depolarized fibres (expts 3, 4).

Table 1 Uptake of dihydropyridines by frog sartorius muscles

Expt	Control	Test	Test/control
1	High K	Ca-free	1.15 ± 0.06
2	Ringer's	Adrenaline, 10 μ M	0.91 ± 0.17
3	Ringer's	³ H-nitrendipine	0.22 ± 0.03
4	High K	³ H-nitrendipine	0.37 ± 0.02

Throughout, control and test conditions were compared on paired muscles (five muscle pairs per experiment) and specific uptake was measured as described in Fig. 2 legend. Control: Incubation at final concentrations of 1.5–1.6 nM ³H-PN-200/110 (Ringer's) or of 2.4–3.4 nM (high-K solution; composition as in Fig. 2). Expt 1, the incubation fluid was Ca-free (40 mM tetraethylammonium (TEA⁺)-EGTA, 65 mM TEA⁺-aspartate, 5 mM TEA⁺-MOPS, pH 7.0) and the muscles were chopped into 1-mm pieces to keep them depolarized and to allow this solution to enter the cytoplasm. Expt 2, as the control, but with 10 μ M adrenaline added to the incubation fluid. Expts 3, 4, as the control, but ³H-nitrendipine (NEN) was used as label instead of ³H-PN-200/110. The final concentrations of ³H-nitrendipine (1.49 ± 0.06 nM in expt 3, and 3.37 ± 0.23 nM in expt 4) were larger than those of ³H-PN-200/110; therefore, the uptake of ³H-PN-200/110 was corrected (using $K_D = 1.0$ nM) to apply to the final concentrations obtained with ³H-nitrendipine.

It is of interest to consider the number of binding sites observed here. With 80 pmol g⁻¹, there are four times more dihydropyridine sites than there are Na channels¹⁵; with a transverse tubule membrane area of 2,200 cm² per m of fibre volume¹⁶, there are ~230 sites per μ m² of that membrane. 80 pmol g⁻¹ correspond to $\sim 5 \times 10^{13}$ sites per ml fibre volume, each representing a protein complex of relative molecular mass ~210,000 (ref. 17). This number of sites is large compared with the observed Ca current. At ~-20 mV, where Ca current is largest (Fig. 3a), its average value is 102 ± 9 mA ml⁻¹ ($n = 6$). This corresponds to 2 fA per site, much less than the current through single Ca channels observed at similar potentials and external Ca²⁺ concentration (~0.1 pA in chromaffin cells¹⁸ or 0.07 pA in cardiac myocytes¹⁹). Evidently, there are 35 to 50 times more dihydropyridine binding sites than there are voltage-dependent Ca channels that can be activated to pass current. This difficulty is not removed by considering that, at the normal negative resting potential, most dihydropyridine receptors have an undetectably low or no affinity for PN-200/110. Even recep-

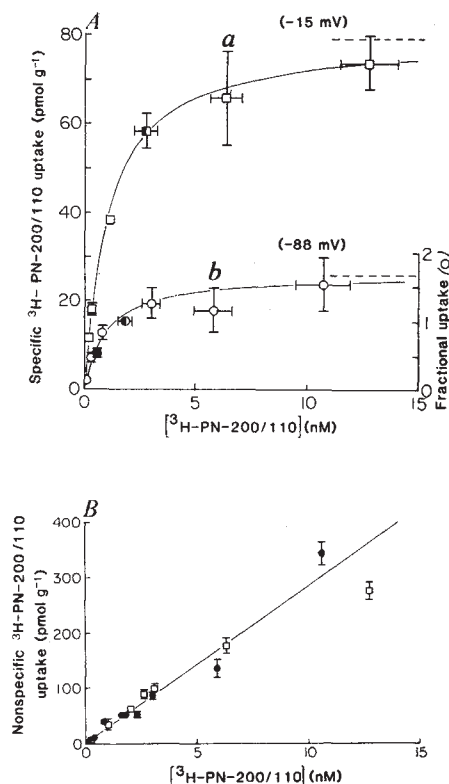


Fig. 2 A, Saturable uptake of ³H-PN-200/110 determined from the drug contents and concentrations before and 2 h after adding an excess of unlabelled nitrendipine (see Fig. 1b and text). B, Non-saturable uptake against concentration in Ringer's (filled symbols) and high-K solution (squares).

Methods. Paired sartorius muscles were used; the first was incubated at an approximately fixed final reference concentration (1.9 nM in curve a; 2.7 nM in b), the second at various test concentrations. Specific uptakes (calculated from the displacement of label by nitrendipine) at the test concentrations were normalized to that at the reference concentration (the ordinate on the right of curve a is included as an example). Each data point represents measurements from at least five muscles; s.e.m. (uptakes) or s.d. (final concentrations) are indicated. Curve a, in Ringer's solution as in Fig. 1b; curve b, in a high-K solution with the same Ca²⁺ activity as Ringer's solution (40 mM K₂SO₄, 5 mM CaSO₄, 2 mM Tris-maleate, 113 M sucrose, pH 7.2). On first contact with the K-rich solution, muscles went through a transient contraction, during which time they were kept pinned out in a dish to prevent shortening. We rejected measurements from muscles that showed signs of contracture at the end of the experiment. To measure the membrane potentials in Ringer's, a muscle was incubated for 5 h and transferred to a dry recording chamber, then the incubate was poured over it. The average resting potential in this muscle was 87.5 ± 0.6 mV ($n = 15$). The membrane potential in the high-K solution is ~-15 mV. Curves a and b are least-squares fits of the function $B_s = B_{max} \times [\text{label}] / ([\text{label}] + K_D)$. For least-squares fitting, uptakes and final concentrations were entered into the computer for each muscle; for the uptake, the ratio of test/reference was entered for the test muscle and a value of 1.0 for the reference muscle. Values for B_{max} giving the best least-squares fit (dashed lines) were 1.71 ± 0.04 times (Ringer's) or 1.35 ± 0.04 times (high-K solution) the reference uptake. Half-filled symbols represent uptakes (in pmol g⁻¹) at the reference concentrations; their values were 15.4 ± 0.7 pmol g⁻¹ ($n = 33$) in Ringer's and 58.7 ± 3.8 pmol g⁻¹ ($n = 25$) in the high-K solution. These values define the ordinate on the left. At low concentrations, saturable and nonsaturable uptakes can be obtained more directly by comparing paired sartorius muscles incubated for 5 h in high-K solution. One muscle was incubated without, and the contralateral muscle with, excess (700 nM) nitrendipine. Digesting the muscles and then counting radioactivity showed that, at 0.71-nM label, the uptake with nitrendipine was 24.4 ± 0.5 pmol g⁻¹ less than that without (five muscle pairs). A group of five muscles from the same batch of frogs and processed as described in Fig. 1b took up 26.5 ± 4.1 pmol g⁻¹ at 0.77 nM label, an identical result.

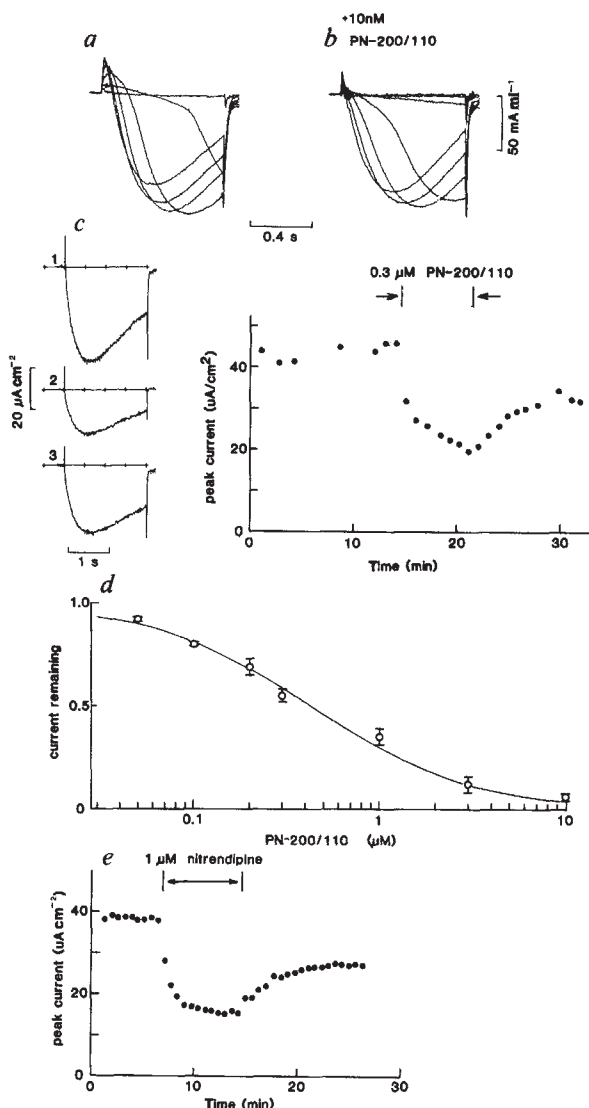


Fig. 3 *a, b*, Ca currents recorded with the three-electrode voltage clamp²² from frog sartorius muscle fibres in a solution containing 107 mM TEA-methylsulphonate, 10 mM Ca-(methylsulphonate)₂, 10 mM MOPS buffer (pH 7.2) and 1 μM tetrodotoxin. ³H-PN-200/110 binding in this solution was 1.40 ± 0.22 times that in Ringer's solution (five muscle pairs). 350 mM sucrose was added to abolish contraction. Holding potential -90 mV; current traces were recorded during steps to between -60 or -50 mV and 0 mV in 10-mV increments. Ordinate calibration: At 23 °C in isotonic solution, the myoplasmic resistivity is calculated to be $150 \Omega \text{ cm}$ (ref. 23); adding 350 mM sucrose at that temperature increased the myoplasmic resistance per unit length 1.49 ± 0.13 -fold ($n = 7$, microelectrode cable analysis). Consequently we used a myoplasmic resistivity of $223 \Omega \text{ cm}$ in equation (4) of ref. 22 to refer currents to the fibre volume in isotonic conditions. *a*, Control; *b*, from the contralateral muscle of the same animal after 6 h incubation in Ringer's +10 nM PN-200/110. The drug was retained continuously also after the incubation and during the recording period. *c*, Left, Ca-current during step depolarizations from the holding potential (-80 mV) to 10 mV before (1), during (2) and after (3) the presence of 300 nM PN-200/110. Right, peak current as a function of time, referred to the surface of a cylinder of the same length and diameter as the piece of fibre from which current was collected. External solution as in *a, b* except that sucrose was omitted. Vaseline-gap voltage clamp⁷ with fibre ends cut in the solution of Table 1, expt 1. Throughout (*a-c*), currents were corrected for leakage and capacitive currents⁷. *d*, Peak current remaining 4–5 min after drug application as a fraction of that recorded immediately before; each point represents data from 3–8 experiments as in *c*. PN-200/110 (a gift of Dr C. E. Eden, Sandoz) was a 1/1 mixture of (+) and (-) stereoisomers. Their individual potencies are not known¹² but in the worst case one isomer would be inactive and the concentrations for the active isomer would be half those indicated. *e*, As *c*, but with nitrendipine.

tors in the low-affinity form are so abundant ($79 - 26 = 53 \text{ pmol g}^{-1}$ in Fig. 2) that at most a few per cent of them can represent functional Ca channels.

Figure 3*b* demonstrates that PN-200/110 has no or only small effects on Ca currents at concentrations high enough to load nearly all high-affinity sites. Although this muscle had been exposed continuously to 10 nM PN-200/110 for the previous 6 h, it retained Ca currents that were as large, or almost as large ($87 \pm 5 \text{ mA ml}^{-1}$, $n = 7$), as in the untreated contralateral muscle from the same animal (102 mA ml^{-1}). The slightly smaller average value with drug may result from the longer time that elapsed between dissection and recording (9–11 h). Hence, the drug can bind without blocking Ca channels, as has been suggested for synaptosomes²⁰.

At higher concentrations, PN-200/110 causes block, as do other 1,4-dihydropyridines. In Fig. 3*c*, 300 nM drug diminished Ca current to ~47% within 6 min, with partial recovery of current after drug removal. Other, similar experiments yielded the dose-response curve in Fig. 3*d*. The curve describes block by one-to-one binding of PN-200/110 to Ca channels ($K_d \sim 430 \text{ nM}$). Prolonged application does not cause much stronger block. In other experiments (not shown), a semitendinosus muscle was incubated in 300 nM PN-200/110 for 5 h; then, with the drug still present, three single fibres were dissected over the next 4 h and currents recorded. Peak currents averaged $27 \pm 5 \mu\text{A}$ per cm^2 of fibre surface membrane. Four fibres from the drug-free contralateral muscle had previously given $58 \pm 9 \mu\text{A cm}^{-2}$. The difference is roughly consistent with Fig. 3*d*. Figure 3*e* shows block of Ca current by nitrendipine; this and two other such experiments suggest that ~700 nM of the drug is required to block half the channels.

Thus, PN-200/110 can bind without causing block, and there are many more high- (and low-) affinity binding sites than there are Ca channels. Both findings are explained most easily if most binding sites for antagonist 1,4-dihydropyridines are not functional Ca channels. In skeletal muscle, they may be unrelated to Ca channels, but this view is contrary to the circumstantial evidence in smooth and cardiac muscle and in pheochromocytoma cells²¹. An alternative speculation is that dihydropyridine receptors represent a large reserve of dormant Ca channels kept for an unknown physiological emergency.

The voltage dependence of dihydropyridine binding described here may be compared with models proposed for cardiac muscle¹¹, where low- and high-affinity binding sites are thought to represent normal and inactivated Ca channels, respectively, that are in voltage-dependent equilibrium. In such a model, Ca channels should bind drug with a single apparent K_d (given in ref. 11) that diminishes in a graded fashion as the membrane potential is made more positive and the fraction of inactivated channels increases. The prediction is a competitive inhibition of drug binding by negative potentials, with a change of K_d but not of B_{max} . Figure 2 shows the opposite. In skeletal muscle, negative potentials evidently inhibit binding in a non-competitive fashion.

We thank Laura Bolles for technical help, Dr B. M. Curtis for suggesting the use of PN-200/110, Dr W. A. Catterall for helpful discussions and Drs E. Neher and W. Stuhmer for comments on the manuscript. L.M.S. held a postdoctoral fellowship of the American Heart Association. This work was supported by DHHS grant AM17803.

Received 17 September 1984; accepted 6 February 1985.

1. Fosset, M., Jaimovich, E., Delpont, E. & Lazdunski, M. *J. Biol. Chem.* **258**, 6086–6092 (1983).
2. Glossman, H. & Ferry, D. R. *Archs Pharmac.* **323**, 279–291 (1983).
3. Curtis, B. M. & Catterall, W. A. *Biochemistry* **23**, 2113–2118 (1984).
4. Gould, R. J., Murphy, K. M. & Snyder, S. H. *Molec. Pharmac.* **25**, 235–241 (1984).
5. Bolger, G. T. et al. *Biochem. biophys. Res. Commun.* **104**, 1604–1609 (1982).
6. Fleckstein, A. *Calcium Antagonism in Heart and Smooth Muscle* (Wiley, New York, 1983).
7. Almers, W., Fink, R. & Palade, P. T. *J. Physiol., Lond.* **312**, 117–207 (1981).
8. Kass, R. S. *J. Pharmac. exp. Ther.* **223**, 446–456 (1982).
9. Lee, K. S. & Tsien, R. W. *Nature* **302**, 790–794 (1983).
10. Almers, W. & McCleskey, E. W. *J. Physiol., Lond.* **353**, 586–608 (1984).
11. Bean, B. P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6388–6392 (1984).
12. Muller-Schweinitzer, E. & Neumann, P. J. *J. cereb. Blood Flow Metab.* **3**, 356–361 (1983).
13. Ferry, D. R., Goll, A. & Glossmann, H. *Archs Pharmac.* **323**, 276–277 (1983).
14. Reuter, H. & Scholz, H. *J. Physiol., Lond.* **264**, 17–47 (1977).
15. Almers, W. & Levinson, S. R. *J. Physiol., Lond.* **247**, 483–509 (1975).

16. Mobley, B. A. & Eisenberg, B. R. *J. gen. Physiol.* **66**, 31-46 (1975).
17. Norman, R. I., Borsotto, M., Fosset, M. & Lazdunski, M. *Biochem. biophys. Res. Commun.* **111**, 878-883 (1983).
18. Fenwick, E. M., Marty, A. & Neher, E. *J. Physiol., Lond.* **331**, 599-635 (1982).
19. Hess, P. & Tsien, R. W. *Nature* **309**, 453-456 (1984).
20. Nachshen, D. A. & Blaustein, M. P. *Molec. Pharmacol.* **16**, 579-586 (1979).
21. Toll, L. *J. biol. Chem.* **257**, 13189-13192 (1982).
22. Adrian, R. H., Chandler, W. K. & Hodgkin, A. L. *J. Physiol., Lond.* **208**, 607-644 (1970).
23. Hodgkin, A. L. & Nakajima, S. *J. Physiol., Lond.* **221**, 105-120 (1972).

Central nervous system and peripheral nerve growth factor provide trophic support critical to mature sensory neuronal survival

Eugene M. Johnson Jr & Henry K. Yip

Department of Pharmacology, Washington University School of Medicine, 660 South Euclid Avenue, St Louis, Missouri 63110, USA

Primary sensory neurones in cranial and dorsal root ganglia (DRG) of adult animals are generally thought to be maintained through connections with their peripheral (but not central) targets by trophic factor(s) other than nerve growth factor (NGF)¹. Damage to the peripheral process of sensory neurones results in a dramatic response or even death of the neurones, whereas axotomy (cutting) of the central process does not initiate profound reaction in these neurones. The development and maintenance of neurones are highly dependent on a supply of trophic agents produced by targets and retrogradely transported via the peripheral process to the cell body². NGF deprivation in fetal rodents produced either by exogenously administered antibodies or by those of maternal origin, results in death of DRG²⁻⁵ and of some cranial sensory neurones⁶. However, as chronic NGF deprivation in neonatal or adult rodents produces little or no cell death⁷⁻¹⁰, it has been assumed that some other trophic factor(s) derived from the peripheral target sustains sensory neurones in postnatal life. By inducing NGF deprivation by autoimmunizing guinea pigs with mouse NGF and/or by cutting the central root (process) of a DRG, we demonstrate here that under certain conditions DRG neurones require NGF and centrally derived trophic support. Our results indicate that sensory neurones are maintained by the trophic support provided by both peripheral and central targets. This support is mediated by NGF and other as yet unidentified trophic factors. The relative importance of the two target fields and NGF compared with other trophic factors changes during development.

Chronic peripheral NGF deprivation was produced in young adult Hartley guinea pigs (Charles River) by immunizing with 2.5S NGF as described previously^{4,8}. Age-matched control animals were not immunized. Two months after the immunization of the experimental group, surgical lesions to the central and/or peripheral processes of sensory neurones were performed as described in Fig. 1 legend. Three months after surgery, the effect of the single or combined lesions on neuronal number in L₆ DRG of nonimmunized control guinea pigs is shown in Table 1. Sciatic nerve crush injury results in a loss of ~20% of L₆ DRG neurones. As 80-85% of L₆ DRG neurones project down the sciatic nerve⁸, 25% of the axotomized neurones die. In contrast, cutting of the dorsal root does not result in any cell loss, consistent with many reports and the data of Hinsey *et al.*¹¹, who found no decrease in neuronal number of S₂ lumbar ganglion in adult cat after dorsal root section. The data in Table 1 also show that combined lesions of the central and peripheral processes do not result in a statistically significant additional cell death, consistent with the observations of Hare and Hinsey¹² who observed that lesion of the central process did not increase the severity of the neuronal response to crushing of the sciatic nerve in cats killed 6-35 days after surgery.

The consensus that the central process has little or no role in providing trophic support of the sensory neurone clearly is not the case in immature animals. We have shown¹³ in newborn rats that a loss of 50% of L₅ DRG neurones is produced by either

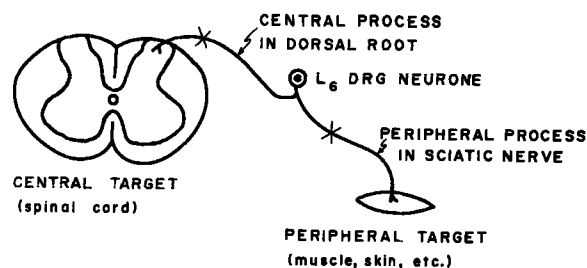


Fig. 1 Anatomy of the L₆ spinal sensory neurone showing the terminal areas of its central (spinal cord) and peripheral processes (for example, muscle). Crosses indicate lesion sites. Two months after immunization of the experimental group, animals were anaesthetized with chloral hydrate (350 mg per kg). Animals receiving a central lesion underwent an L₆ dorsal rhizotomy following exposure by a thoraco-lumbar laminectomy. The L₆ central process was sectioned near its entry to the spinal cord. Animals receiving a peripheral lesion underwent axotomy of their sciatic nerves. The nerve was exposed in mid thigh and crushed for 40 s with a pair of jeweller's forceps at the tendon of the *obturator internus*. Animals were prepared with either a central or peripheral lesion alone or a combination of the lesions. Corresponding sham-operated control animals were prepared. Immunized animals were boosted with NGF monthly to maintain their titre of anti-NGF, and 90-100 days after lesion, they were perfused transcardially with saline and then with 10% formal saline. The sixth lumbar ganglia were removed bilaterally and immersed in 10% formalin and subsequently embedded in paraffin using standard histological techniques; 8-μm serial sections of the ganglia were stained with toluidine blue. The slides were coded so that their identity was unknown to the examiner. In every tenth section, neurones with a nucleolus were counted and the raw counts corrected¹⁷.

a central or a peripheral lesion and that the combined lesion results in the loss of 70% of the neurones. Cell loss produced by either the single or combined lesion can be prevented by concomitant treatment with systemic NGF. Also, ¹²⁵I-labelled NGF can be efficiently and specifically retrogradely transported from the spinal cord to the DRG^{13,14}. These experiments¹³ demonstrate that the central process, as well as the peripheral process, has a critical role in providing trophic support to the developing sensory neurones in newborn rats and suggest that the trophic molecule provided by the central process is NGF.

Peripheral NGF deprivation by chronic exposure to anti-NGF in adults does not produce DRG neuronal death nor does it increase neuronal death in peripheral nerve crush⁸. It is clear, however, that even in adults, NGF deprivation has an adverse effect on DRG neurones. Chronic NGF deprivation in adults results in decreased levels of substance P⁷ and a modest atrophy of neurones in DRG⁸. Thus, in contrast to the dependence of immature sensory neurones on peripheral NGF for survival, the mature neurone atrophies but does not die as a result of peripheral NGF deprivation. Similarly, the loss of trophic support from the central process results in cell death of immature DRG neurones; however, this loss of trophic support in adults is insufficient to cause cell death. We speculate that the conclusion that mature DRG neurones do not require NGF for survival may be incorrect because it ignores the possibility that

Table 1 Neuronal numbers in L₆ DRG after lesioning of the central and/or peripheral processes in normal guinea pigs

Group	Lesion		Cell no. in L ₆ DRG	% Control
	Central	Peripheral		
A	-	-	17,897 ± 711 (9)	100
B	+	-	17,439 ± 695 (4)	97
C	-	+	14,304 ± 1,184 (5)	79*
D	+	+	12,777 ± 847 (3)	71*

The data were analysed using two-way analysis of variance including Satterthwaite's correction of degrees of freedom resulting from unequal group size¹⁶.

* Statistically different from control (group A); *P* < 0.02.

Table 2 Neuronal number in the L₆ DRG after lesioning of the central and/or peripheral processes in NGF-immunized guinea pigs

Group	Lesion		Cell no. in L ₆ DRG	% Control
	Central	Peripheral		
A	—	—	18,113 ± 815 (7)	100*
B	+	—	13,589 ± 996 (3)	75
C	—	+	14,695 ± 726 (3)	81
D	+	+	10,641 ± 711 (4)	59*

* Statistically different from all other groups (maximum $P < 0.054$).

NGF may be taken up by central terminals in the spinal cord and retrogradely transported by the central process to the DRG. Antibodies to NGF presumably would not cross the blood-brain barrier and would not prevent access of the DRG neurone to central nervous system (CNS)-derived NGF (or other trophic factor).

The data in Table 2 show the effects on sensory neurones of lesioning either the central or peripheral process or both in a group of animals chronically deprived of NGF. The peripheral lesion alone results in ~20% loss of DRG neurones⁸. However, in the NGF-deprived animal, cutting only the central process resulted in a 25% loss. In anti-NGF-producing animals subjected to a peripheral crush and a central lesion, a further neuronal loss of 27% (14,695–10,641) is observed. Thus, when the trophic support of peripherally derived NGF is removed, at least some populations of sensory neurones become dependent on centrally derived trophic support for survival. Conversely, when deprived of trophic support through the central process, at least a proportion of sensory neurones requires peripheral NGF for survival.

These results suggest that survival of sensory neurones in the adult animal depends on trophic support from both the central and peripheral targets. As peripheral NGF deprivation produces atrophy but not cell death of DRG neurones, whereas crushing the nerve does produce (~25%) cell death, we suggest that NGF is but one of possibly several peripherally provided molecules that trophically support DRG neurones. As NGF deprivation plus central lesion produces an amount of cell death comparable with that caused by peripheral crush lesion, the total trophic support provided by the central process is equivalent to the non-NGF component of peripheral trophic

support. Thus, the loss of any one of these three trophic components (peripheral NGF, peripheral non-NGF, centrally derived factor) produces atrophy but not death. Removal of any two of these components results in loss of ~20% of cells. Loss of all three components would therefore be expected to result in the death of even more cells (Table 2). The data in Table 1 suggest that central process lesion may increase the severity of a peripheral crush in control animals, but the decrease in neuronal number is not statistically significant. Thus, we propose that the trophic support of a DRG neurone in the adult animal is represented by the sum of the trophic effects of NGF and other unknown factors derived from the periphery and a factor or factors derived from the CNS, one of which may be NGF. Based on our limited data, we cannot provide a complete explanation of this hypothesis. Interpretation is complicated by the fact that the DRG is composed of heterogeneous cell types. We do not know, for example, whether the 20% of cells that die after peripheral lesion are the same cells that die after peripheral NGF deprivation combined with a central lesion. In addition, the experiment used a peripheral crush lesion rather than a cut; thus, regeneration of the distal processes was possible and probably rather efficient⁸. Greater cell loss would be expected if the sciatic nerve were cut rather than crushed. Similarly, 15–20% of neurones in L₆ DRG do not project down the sciatic nerve and some neurones may reach the CNS via the ventral root, rather than the dorsal root¹⁵.

Irrespective of the complications in providing a precise accounting of the cell death, the fact that trophic support for sensory neurones is provided via their central processes seems to explain the apparent lack of dependency of mature sensory neurones on NGF for survival, deduced by experiments using anti-NGF. The data suggest that survival of sensory neurones is dependent on relative trophic support provided by both the central and peripheral processes, although the quantitative support by the peripheral process is overall more important in adult animals. Because in newborn rats cell losses after lesioning of the central or peripheral process are the same¹³, the results indicate that the relative importance of the trophic support provided by the central and peripheral processes changes as a function of age.

This work was supported by NIH grants NS18071, HL20604 and NS07129. We thank Ms Patricia Osborne and Ms Patricia Lampe for assistance and Dr Keith Rich for advice.

Received 24 September 1984; accepted 12 February 1985.

- Lieberman, A. R. *Int. Rev. Neurobiol.* **14**, 49–124 (1971).
- Hamburger, V. & Oppenheim, R. W. *Neurosci. Comment.* **1**, 39–55 (1982).
- Gorin, P. D. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5382–5386 (1979).
- Johnson, E. M., Gorin, P. D., Brandeis, L. D. & Pearson, J. *Science* **210**, 916–918 (1980).
- Aloe, L., Cozzani, C., Calissano, P. & Levi-Montalcini, R. *Nature* **291**, 413–415 (1981).
- Pearson, J., Johnson, E. M. & Brandeis, L. *Dev. Biol.* **96**, 32–36 (1983).
- Schwartz, J. P., Pearson, J. & Johnson, E. M. *Brain Res.* **244**, 378–381 (1982).
- Rich, K. M., Yip, H. K., Osborne, P. A., Schmidt, R. E. & Johnson, E. M. *J. comp. Neurol.* **230**, 110–118 (1984).
- Levi-Montalcini, R. & Angeletti, P. U. *Physiol. Rev.* **48**, 534–569 (1968).
- Thoenen, H. & Barde, Y. A. *Physiol. Rev.* **60**, 1284–1335 (1980).
- Hinsey, J. C., Krupp, M. A. & Lhamon, W. T. *J. comp. Neurol.* **67**, 205–214 (1937).
- Hare, W. K. & Hinsey, J. C. *J. comp. Neurol.* **73**, 489–502 (1940).
- Yip, H. K. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6245–6249 (1984).
- Richardson, P. M. & Riopelle, R. J. *J. Neurosci.* **4**, 1683–1689 (1984).
- Coggeshall, R. E. *Neurosurgery* **4**, 443–448 (1979).
- Snedecor, G. W. in *Statistical Methods* (Iowa University Press, 1957).
- Konigsmark, B. E. in *Contemporary Research Methods in Neuroanatomy* (eds Nauta, W. H. & Ebbesson, S. O. E.) 315–340 (Springer, New York, 1970).

Erratum

Lower mantle heterogeneity, dynamic topography and the geoid

B. H. Hager, R. W. Clayton, M. A. Richards, R. P. Comer & A. M. Dziewonski
Nature **313**, 541–545 (1985).

IN this article Table 1 was omitted. It should read:

Table 1 Correlation coefficients for degrees 2–6 for model geoids predicted using two models of lower mantle heterogeneity (Dz and CC), convolved with dynamic response functions for Earth model U10

Degree	Model geoid CC-U10					Model geoid Dz-U10				
	2	3	4	5	6	2	3	4	5	6
Observed geoid	0.84	0.76	0.34	−0.17	−0.02	0.68	0.74	−0.24	0.24	0.17
Residual geoid	0.95	0.77	−0.25	−0.35	−0.33	0.82	0.79	−0.39	0.27	−0.03
Model geoid Dz-U10	0.79	0.66	0.01	0.12	0.31					

The models are correlated against the observed geoid, a residual geoid with slab effects removed¹⁰, and each other.

Greek archaeomagnitudes

RELIABLE determination of the ancient geomagnetic field intensity is difficult; any new approach, such as that of Walton¹, is to be welcomed. Although the method is new and complex, it does not follow that it necessarily gives reliable answers, particularly as its validity rests, on various theoretical assumptions about the behaviour of the magnetic carriers. We consider that Walton's dismissal of the linearity test in the commonly used Thellier technique is premature and insecurely founded. Certainly, this technique always needs to be carried out meticulously, as emphasized by Thellier, otherwise erroneous results are obtained that can, if taken at face value, lead to the supposition of rather exotic behaviour by the geomagnetic intensity in the past. Hence, Walton's finding that his earlier determinations using the Thellier technique were unreliable does not mean that all Thellier results are suspect.

In any linearity test the critical question is 'How linear is linear?' and in the examples referred to by Walton, it may well be that the criteria used were not sufficiently strict. On the other hand, we have confirmed the effectiveness of reliability criteria, where they are stringently applied, by re-examining our recent Thellier results² for the second millennium BC; we have remeasured several of our samples using the coercivity monitoring technique developed by Shaw³ and find that the agreement is good. In this method, the coercivity spectrum of the natural remanent magnetization (NRM, the ancient magnetization) is compared with that of the thermoremanent magnetization (TRM, the magnetization acquired on cooling after the laboratory heating); changes in shape indicate that the laboratory heating has caused mineral alteration and there is risk of error. Also, the coercivity spectrum of an acquired remanent magnetization (ARM, magnetization

acquired from a weak steady field in the presence of a strong alternating field) given before heating is compared with that of an ARM given after heating; as well as spectrum shape, this monitors any increase or decrease in the magnetic carriers in each coercivity interval, thus providing a built-in quantitative measure of the magnetic alteration. In the basic Shaw technique the ancient intensity is obtained from the NRM/TRM ratio evaluated from coercivity intervals for which there is no change in ARM strength. In later developments^{4,5} it has been shown that, even if alteration occurs, the ARM ratio can be used to correct the NRM/TRM ratio, at least for coercivity intervals >100 mT.

We have now used the Shaw technique, or one of its variants, on eleven samples from several of the sites in ref. 2 (see Table 1). There is agreement with the Thellier values to within 15% for nine of the samples; seven of these are within 10% with the Shaw values on average 5% lower than the Thellier values. The two remaining samples gave inconsistent results for different coercivity intervals, although always within $\pm 25\%$ of the Thellier value. These are our first results using the Shaw technique and, because there is scope for improvement in our facilities, we consider that good confirmation of the validity of the Thellier results has been obtained. Also, we have used the ARM technique on subsidiary cores (19 samples), with intervening heatings to various of the temperature steps used in the Thellier determination. In this way we can determine whether or not alteration has set in by the upper temperature limit of accepted Thellier data; in no case has the ARM change at that limit been >6%. These tests justify our confidence in the results of ref. 2.

Finally, regarding the alteration test of comparison between two measurements made in zero field, suggested by Walton for the Thellier method, we note that his procedure monitors only those carriers having blocking temperatures greater than

the magnetization temperature; on the other hand, the in-field measurement is also affected by carriers having lower blocking temperatures. Our experience of comparing demagnetization curves before and after heating indicates that alteration can occur in some blocking temperature ranges without affecting others. A much better test is to compare two in-field measurements, but tests of this type can never give conclusive proof that alteration has not occurred because of the probability that it occurs rapidly and is effectively complete during the first stage of a given temperature step. These comments apply equally to the Walton method.

M. J. AITKEN
A. L. ALLSOP
G. D. BUSSELL
M. B. WINTER

Research Laboratory for Archaeology,
Oxford University,
6 Keble Road,
Oxford OX1 3QJ, UK

1. Walton, D. *Nature* **310**, 740-743 (1984).
2. Aitken, M. J., Allsop, A. L., Bussell, G. D. & Winter, M. B. *Nature* **310**, 305-306 (1984).
3. Shaw, J. *Geophys. J. R. astr. Soc.* **39**, 133-141 (1974).
4. Kono, M. *Geophys. J. R. astr. Soc.* **54**, 241-261 (1978).
5. Rolph, T. C. & Shaw, J. *Geophys. J. R. astr. Soc.* **80**, 773-781 (1985).

WALTON REPLIES—There is quite a clear disagreement between recent results for Greece¹ and those of Aitken *et al.*²

I have published two sets of archaeomagnetic intensity measurements for Greece. The original data³ also show the increase in intensity that Aitken *et al.* are proposing to use for archaeomagnetic dating. Unfortunately, the old data in ref. 3 are wrong, which prompted the study in ref. 1. This later study showed that the error was caused by mineral alteration at the higher measurement temperatures. A few samples that showed no evidence of alteration did not show an increase in intensity. It was possible to correct for a small degree of alteration; if the effects of alteration were retained, the increase in intensity referred to above was present in the high temperature points but not at low temperatures. On correction, all the points at all temperatures were consistent, within the experimental accuracy, and the increase disappeared.

The experiments reported in ref. 3 used a technique that did not differ substantially from that used by Aitken *et al.*², but Aitken *et al.*² express reservations about the technique used in ref. 1. It is important, however, to recognize that in all three methods the ancient intensity is obtained by the same method, thermal demagnetization and magnetization, differing only in the test used for alteration.

More specifically, the method used by Aitken *et al.*², and initially by myself,

Table 1 Comparison of Shaw and Thellier techniques

Presumed date BC	Sample location	Intensity ratio (Thellier)	Ancient intensity (μT)	
			Thellier	Shaw
1550-1500	Apheq, Israel	1.06	45	43
1550-1500	Apheq, Israel	1.15	49	48
1504-1492	Karnak, Egypt	1.17	45	40
1250-1150	Karnak, Egypt	1.51	58	57
1250-1150	Karnak, Egypt	1.53	59	57
1150-1075	Kition, Cyprus	1.62	72	71
1100-950	Karnak, Egypt	1.52	59	51
1100-950	Karnak, Egypt	1.52	59	61
1050-1000	Kouklia, Cyprus	1.53	68	68

The Thellier results are from ref. 2; the intensity ratio quoted is F_A/F_D , where F_A is the ancient intensity and F_D is the intensity at the site attributable to an axial centred geomagnetic dipole of present-day strength (8×10^{22} A m²).

method 1, compares the moment lost after demagnetizing for a fixed time at a given temperature, with the moment gained in a known field after holding for the same time at the same temperature. My new method, method 2 (see ref. 1), also compares the moment lost in zero field on holding the sample for a fixed time at a temperature with the moment gained in the same time period in the laboratory field. In method 2, because the moment of the sample is being monitored continuously, several readings can be made while the sample is hot and, to improve the statistical accuracy, the slope of the best straight line through the points is used instead of simply taking the difference between the first and last reading. Thus, the 'theoretical assumptions' on which our methods depend are, up to this point, identical.

The two methods do differ in the means they use to detect alteration, as discussed in refs 1,2. The important point is that alteration was detected using method 2, but was not detected using the linearity test of ref. 2.

In the event that alteration is complete by the end of the first heating in method 2, it would still be detected as a sudden change in the value of the ancient intensity at that temperature. Method 2, in other words, also incorporates a 'linearity test'.

In any case, the point is not that alteration has been missed but that it has been detected. Almost all the material between 1200 and 300 BC showed evidence of alteration at temperatures below 350 °C. If the altered points were used, high values were obtained for the ancient field in good agreement with refs 2 and 3. If points with no evidence of alteration were used, the increase that Aitken *et al.* report, forming the basis for their dating method, disappears.

It will be useful to compare these results with those obtained using the Shaw method. At present Aitken *et al.* report nine values. Of these, three are dated 1500 BC where there is no disagreement. In assessing the remaining six it is important to consider their internal consistency: results between 1250 and 950 BC are between 58 and 72 μ T with the Thellier and 51 and 71 μ T with the Shaw method. In the same period, nine lie between 46 and 54 μ T. With so few results the difference is hardly significant.

D. WALTON

Physics Department,
McMaster University,
1280 Main Street West,
Hamilton, Ontario,
Canada L8S 4M1

Bullard Laboratories,
Department of Earth Sciences,
University of Cambridge,
Madingley Rise,
Madingley Road,
Cambridge CB3 0EZ, UK

1. Walton, D. *Nature* 277, 643-644 (1979).
2. Aitken, M. J., Allsop, A. L., Russell, G. D. & Winter, M. *Nature* 310, 306-308 (1984).
3. Walton, D. *Nature* 310, 740-743 (1984).

A Grenville Sm/Nd age for the Glenelg eclogite

SANDERS *et al.*¹ present data indicating that eclogites within the eastern Lewisian inlier at Glenelg equilibrated at ~1,100 Myr BP at temperatures and pressures of ~700 °C and >12 kbar. On the basis that the overlying Moine cover shows no sign of having exceeded conditions for the lower amphibolite facies, they conclude that the Moine was deposited unconformably on a deeply eroded Lewisian basement following Grenville age metamorphism, a view at variance with much recent work which regards the Moine as a series of late mid-Proterozoic metasediments that underwent widespread metamorphism and deformation at ~1,100-1,000 Myr BP²⁻⁵. Whilst the data produced by Sanders *et al.*¹ is not in dispute, their conclusions relating to the conditions of metamorphism of the eclogites and thus the age of sedimentation of the Moine may be questioned.

We consider the suggestion of Sanders *et al.*¹ that the eclogites were metamorphosed at pressures of >12 kbar questionable on four counts. (1) Mineral parageneses and tabulated phase chemistries are not those of an eclogite in the strict sense but of a medium- or high-pressure granulite. Consequently, deriving a pressure by analogy with Green and Ringwood's⁶ eclogite field is inadvisable. More realistic pressures of as low as 6 or 7 kbar could be derived from their granulite field. (2) Although Sanders⁷ implies that he has corroborative evidence from the garnet/plagioclase/kyanite/quartz geobarometer, this is as yet unpublished and, in view of the problems inherent in modelling plagioclase and garnet activities, may not be conclusive. (3) We have examined staurolite/garnet/kyanite/biotite/quartz-bearing gneisses from the eastern Lewisian in which, on textural criteria, we are confident that all coexisting phases are of Grenville rather than Caledonian age. At temperatures of 650-700 °C, staurolite and quartz will not be stable at pressures in excess of 8 kbar^{8,9}. (4) It has been repeatedly shown that garnet-pyroxenite assemblages comparable to those described from Glenelg are stable at relatively low pressures in rocks undergoing dry metamorphism at amphibolite facies conditions¹⁰⁻¹². These pressures could well be consistent with both the pressure stability limits of staurolite and quartz and pressure estimates of 6-8 kbar obtained from mineral assemblages coeval with the early pre-Caledonian metamorphism within the Morar division Moine (D. Barr, personal communication), in which case

there may be no significant difference between the metamorphic conditions of the Eastern Lewisian and those of the Moine.

In this light a more reasonable explanation is that the ~1,100 Myr BP age¹ dates both the equilibration of the eclogites and the initial amphibolite facies metamorphism of the Moine. Isotope data considered to relate directly to the age of initial metamorphism of the Moine includes Rb/Sr whole-rock ages of ~1,100-1,000 Myr BP on the Morar and Lochailort pelites and the Ardour gneiss¹³⁻¹⁵, which are interpreted by Sanders *et al.*¹ as dating sedimentation and diagenesis rather than metamorphism. Whilst the interpretation of the isotope data relating to the metasediments is equivocal¹⁶, we note that the Ardour gneiss is now generally regarded as a member of an extensive suite of deformed granitic intrusions where emplacement was coeval with amphibolite-facies metamorphism and widespread deformation within the Moine.^{3-5,17} Given that there is no indication of any major structural or metamorphic break of Precambrian age between the south-west Moine and Glenelg¹⁸, it seems difficult to avoid the conclusion that the initial metamorphism of the Moine occurred at ~1,100-1,000 Myr BP.

R. A. STRACHAN

P. J. TRELOAR

Department of Geology and Physical
Sciences,
Oxford Polytechnic,
Headington,
Oxford OX3 0BP, UK

1. Sanders, I. S., van Calsteren, P. W. C. & Hawkesworth, C. J. *Nature* 312, 439-440 (1984).
2. Powell, D., Baird, A. W., Charnley, N. R. & Jordan, P. J. *J. geol. Soc. Lond.* 138, 661-673 (1981).
3. Roberts, A. M. & Harris, A. L. *J. geol. Soc. Lond.* 140, 883-892 (1983).
4. Harris, A. L. in *Geology of Scotland* (ed. Craig, G. Y.) 1-22 (Scottish Academic Press, Edinburgh, 1983).
5. Strachan, R. A. *Scott. J. Geol.* (in the press).
6. Green, D. H. & Ringwood, A. E. *Geochim. cosmochim. Acta* 31, 767-833 (1967).
7. Sanders, I. S. *Spec. Rep. geol. Soc. Lond.* 8, 97-100 (1979).
8. Ganguly, J. *J. Petrol.* 13, 335-365 (1972).
9. Yardley, B. W. D. *Neues Jb. Miner. Mh.*, 127-132 (1981).
10. Turner, F. J. *Metamorphic Petrology* p. 524 (McGraw-Hill, New York, 1980).
11. Bryhni, I., Bollingberg, H. J. & Graff, P. R. *Norsk geol. Tidsskr.* 49, 133-225 (1969).
12. De Wit, M. J. & Strong, D. F. *J. Geol.* 83, 609-627 (1975).
13. Brook, M., Powell, D. & Brewer, M. S. *Nature* 260, 515-517 (1976).
14. Brook, M., Powell, D. & Brewer, M. S. *J. geol. Soc. Lond.* 133, 489-496 (1977).
15. Brewer, M. S., Brook, M. & Powell, D. *Spec. Rep. geol. Soc. Lond.* 8, 129-137 (1979).
16. Van Breemen, O., Halliday, A. N., Johnson, M. R. W. & Bowes, D. R. *Geol. J. spec. Issue* 10, 81-106 (1978).
17. Barr, D., Roberts, A. M., Highton, A. J., Parson, L. M. & Harris, A. L. *J. geol. Soc. Lond.* (in the press).
18. Powell, D. *J. geol. Soc. Lond.* 130, 575-593 (1974).

SANDERS REPLIES—Strachan and Treloar favour contemporaneous (1,100-1,000 Myr BP) metamorphism for both the eclogite-bearing basement and the Moine cover in the Glenelg region. To sustain their view, they suggest that the abnormally high metamorphic pressure

(>12 kbar) inferred for the basement¹ is a serious overestimate.

Although published evidence relating to metamorphic pressure/temperature (P/T) conditions for the Glenelg eclogite may not be extensive, I contend that the rock and mineral analyses reported (Table 1, ref. 1) are more than sufficient to demonstrate a high-pressure origin. The two eclogite samples have tholeiitic chemistry and are plagioclase free; they are therefore true eclogites (rather than high-pressure granulites) and reference to Green and Ringwood's² eclogite field is therefore valid. Moreover, omphacite coexists with quartz in S274 and the mole fraction of jadeite in the omphacite is 0.31 (calculated as $\text{Al(VI)} - \text{Al(IV)}$ in a 4-cation formula). Assuming a disordered pyroxene structure, the omphacite composition implies a pressure >15 kbar at 700°C (ref. 3). The temperature estimate of 700°C is, incidentally, supported by a large body of unpublished analytical data relating to garnet/clinopyroxene thermometry and calcite solvus thermometry.

Against this P/T background, Strachan and Treloar's staurolite/quartz-bearing assemblage from the Glenelg basement may best be regarded as a product of the widespread retrogressive metamorphism that has affected the eclogite and other high-grade rocks at Glenelg⁴. According

to Strachan and Treloar, the staurolite/quartz-bearing assemblage is equivalent in terms of P/T conditions to the Moine. Thus, it would seem that retrogression in the basement correlates with prograde metamorphism in the cover, which is, of course, consistent with the view that the Moines were deposited on an eroded Grenville basement. With regard to the Ardour gneiss, I would argue that the case for intrusion at about 1,000 Myr BP is considerably weakened by the real possibility that the gneiss is a metasediment with a derived Rb/Sr isotope system.

In my view, any claim that the Moines have a Grenville metamorphic history carries with it the implication that they underwent early eclogite-facies metamorphism, at least in the Glenelg area. Hitherto, evidence for such an event has not been reported.

I. S. SANDERS

Department of Geology,
Trinity College,
Dublin 2, Ireland

1. Sanders, I. S., van Calsteren, P. W. C. & Hawkesworth, C. J. *Nature* 312, 439–440 (1984).
2. Green, D. H. & Ringwood, A. E. *Geochim. cosmochim. Acta* 31, 767–833 (1967).
3. Holland, T. J. B. *Contrib. Miner. Petrol.* 82, 214–220 (1983).
4. Ramsay, J. G. *Q. J. geol. Soc. Lond.* 11, 487–523 (1958).

Proper solar wind power estimation and planetary radiometric efficiencies

DESCH and Kaiser¹ have computed the average absolute efficiencies of the radio planets for electromagnetic wave production, restricting their attention to the emissions known to be controlled by solar wind energy inputs to the magnetospheres of the Earth, Jupiter and Saturn. Our purpose here is to offer a word of caution about the methods, assumptions and conclusions of their paper and to suggest that radio emissions for unexplored planets, such as Uranus, may be much different from what they determined.

Their method¹ was to compute the solar wind input power, P_i , to the various magnetospheres by estimating the solar wind kinetic energy flux (ρV^3) at the planetary location and to use the square of the Chapman-Ferraro length ($l_0 = (M^2/4\pi\rho V^2)^{1/6}$) as an estimate of the effective energy 'collection' area of the magnetosphere. Here M is the planetary magnetic moment, ρ is the solar wind mass density (g cm^{-3}) and V is the solar wind speed, so that ρV^3 has units of $\text{erg cm}^{-2} \text{s}^{-1}$. Then, using estimates of the planetary radio emission power, P_r (at a standard distance of 1 AU), the authors computed an efficiency, from the ratio P_r/P_i .

The fault with this method is demon-

strated in Fig. 1. Visual inspection of Fig. 1 shows that the AE index correlates quite well with the ϵ , or P_{VB} , parameter (correlation coefficient ~ 0.7), but the time profile of AE bears little specific relationship to the P_k parameter. Detailed study has shown conclusively^{2–4} that ρV^3 , in the case of the Earth, is not the solar wind parameter that 'drives' auroral disturbances. The results shown in Fig. 1 are quite general. As AE is a measure of auroral activity and can be used³ to estimate auroral particle precipitation and auroral Joule heating, we may conclude with confidence that P_k does not correlate well with these primary forms of terrestrial auroral energy dissipation. Furthermore, as it is well known⁵ and indeed implicit in the argument of Desch and Kaiser, that auroral kilometric radiation (AKR) is well correlated with AE, we may also conclude that P_k is not well correlated with AKR.

The reasoning is thus as follows. The quantity ρV^3 is poorly correlated (on substorm timescales) with VB^2 in the solar wind, but it is VB^2 (or a refinement thereof) which controls substorm (AE) activity at Earth. Because it is clearly substorm activity, and collective plasma processes during substorms, that control the generation of AKR, solar wind estimates of ρV^3 are an improper estimate of P_i as a causal agent of AKR (or P_r), in the case of the Earth. By extension, the computed

radiometric efficiency estimates¹ must be regarded as being predicated on erroneous assumptions.

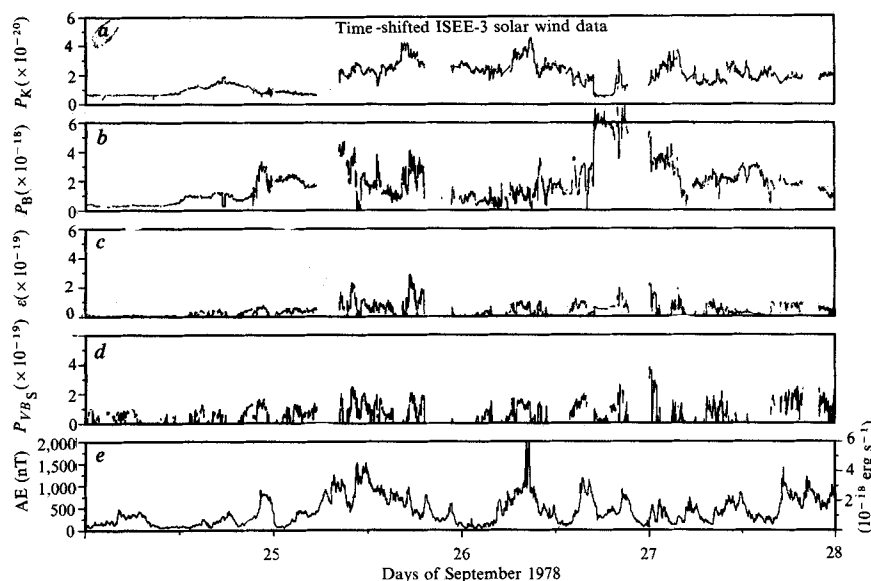
Now Gallagher and D'Angelo have shown a very strong dependence of AKR on the solar wind speed only⁶, but as they used very long-term averages, their results were strongly time aliased. Indeed, if one looks at long-term averages, geomagnetic activity is well correlated with the solar wind speed^{4,7}, but this says almost nothing about the solar wind-magnetosphere coupling mechanism or substorm energy source as substorms occur on timescales of ≤ 1 h.

The criticism raised here might be less significant if ρV^3 were of a similar magnitude as $VB^2/8\pi$; but it is almost always true that ρV^3 is ~ 100 times larger than $VB^2/8\pi$. This is well illustrated in Fig. 1. Thus, as it is the electrostatic⁴ or electromagnetic (Poynting)³ energy in the solar wind which drives auroral activity, the values of efficiency derived by Desch and Kaiser should be ~ 100 times greater in the case of the Earth⁸.

We really do not know how similar or dissimilar the generation of Saturn kilometric radiation (SKR) is to AKR. Furthermore, we are uncertain of the generation mechanism of the hectometre-wavelength (HOM) emission of Jupiter. A further exacerbating feature is that we have little clear evidence as to whether solar wind-magnetosphere coupling proceeds by the same physical processes at Saturn and Jupiter as it does at the Earth. If precisely the same processes were to occur (that is, dayside magnetic reconnection and magnetotail substorm initiation⁸), we would estimate that the SKR and HOM efficiencies should also be ~ 100 times greater than estimated by Desch and Kaiser. On the other hand, if solar wind kinetic energy flux variation were driving the HOM and/or SKR components, then the values of efficiency would be $\sim 5 \times 10^{-4}$ for the Earth, but would be only $\sim 5 \times 10^{-6}$ for Jupiter or Saturn (as estimated by Desch and Kaiser).

In conclusion, based on a proper computation of the energy entering the Earth's magnetosphere, we conclude that terrestrial AKR is produced with a much higher efficiency than computed by Desch and Kaiser. Similarly high efficiencies may or may not apply in the cases of Jupiter and Saturn; this depends very strongly on the physical processes responsible for solar wind-magnetosphere coupling in those cases. Because of such uncertainties, the radiometric estimates made for Uranus¹ must be regarded with caution. In view of the 'pole-on' interaction of Uranus with the solar wind mentioned by Desch and Kaiser, the nature of solar wind coupling may be very different for any possible Uranus magnetosphere. At the very least, readers should appreciate that in the Earth's magnetosphere it is electromagnetic, not particle, energy in the solar wind that drives geomagnetic activity.

Fig. 1 *a-d*, Summaries of various power estimates (in erg s^{-1}) using data taken in September 1978 from the ISEE-3 spacecraft $\sim 240 R_E$ from Earth in the upstream solar wind. These data have been time-shifted to account for propagation time delay to the Earth. *a*, $P_k (= \pi l_0^2 \rho V^3)$ plotted for a several-day interval. *b*, The solar wind magnetic energy power $P_B (= \pi l_0^2 V (B^2/8\pi))$. *c*, *d*, Popular refinements of the magnetic energy power parameter: *c* shows $\varepsilon (= VB^2 \sin^4(\theta/2) (l_0^*)^2)$, where $l_0^* = 7R_E$ and θ is the angle between the Earth's field and the interplanetary field; *d* shows $VB_s (= -B_z$ for negative IMF, and 0 for positive IMF) scaled to the Chapman-Ferrary interaction area. *e*, The Auroral electrojet index, AE (scale to the left). Using empirical relationships^{2,3} it is also possible to make magnetospheric energy dissipation estimates from AE (scale to the right).



This work was performed under the auspices of the US Department of Energy.

D. N. BAKER
L. F. BARGATZE

Los Alamos National Laboratory,
Los Alamos,
New Mexico 87545, USA

- Desch, M. D. & Kaiser, M. L. *Nature* **310**, 755-757 (1984).
- Baker, D. N. *et al.* *J. geophys. Res.* **88**, 6230-6242 (1983).
- Akasofu, S.-I. *Space Sci. Rev.* **28**, 121-190 (1981).
- Baker, D. N., Hones, E. W., Payne, J. B. & Feldman, W. C. *Geophys. Res. Lett.* **8**, 179-182 (1981).
- Voots, G. R., Gurnett, D. A. & Akasofu, S.-I. *J. geophys. Res.* **82**, 2259-2266 (1977).
- Gallagher, D. L. & D'Angelo, N. *Geophys. Res. Lett.* **8**, 1087-1089 (1981).
- Crooker, N. U., Feynman, J. & J. T. Gosling *J. geophys. Res.* **82**, 1933-1937 (1977).
- Baker, D. N. *Am. Inst. Phys. Conf. Proc. Ser. Magnetospheric Phenomena in Astrophysics* (eds Epstein, R., and Feldman, W.) (AIP, New York, in the press)

DESCH AND KAISER REPLY—Baker and Bargatze have argued that high time resolution comparisons of auroral electrojet (AE) with the solar wind input lead to a much higher efficiency for the conversion of solar wind power into auroral kilometric radiation (AKR) than that estimated by us¹. Therefore, if different efficiencies apply at Jupiter and Saturn, where physical processes are not as well understood, then our 'radiometric Bode's law' cannot be reliably extrapolated to the case of Uranus.

We agree that high time resolution studies have led to a better understanding of solar wind-magnetosphere coupling processes at the Earth than presently exists at either Jupiter or Saturn, but it is not clear that this understanding extends to solar wind-AKR coupling, nor is it

apparent that conclusions drawn from longer-term averaging of the data are necessarily improper. Regarding the first point, Voots *et al.*² actually noted a rather small (51%) degree of correlation between AE and AKR, attributing this, in part, to the 'poor short timescale (< 1 h) correlation' that they observed. Thus, although the relationship between substorm activity (as indicated by AE) and the solar wind may be well understood, that between AKR power and the solar wind is not.

We used the solar wind kinetic energy flux, ρV^3 (where ρ is the solar wind mass density and V the solar wind speed), to characterize the energy input to each magnetosphere, because the radio emissions of the known radio planets show by far the best correlation with solar wind plasma properties, such as ρ or V , and little or no correlation with magnetic properties such as VB^2 (refs 3-5). For example, there is almost a one-to-one correspondence between the radio emission output of the AKR and the solar wind speed³ and between the radio output of Saturn and the solar wind density⁴. Each of the studies of Desch and Kaiser has in common the consistent use of one-rotation-averaged solar wind and radio emission quantities. Certainly details of the interaction mechanism are lost when such long-term averages are used, but some coarse information about the interaction probably remains.

Our radiometric Bode's law is intended as a heuristic demonstration of one property of radio-emitting magnetospheres, namely, their gross radio efficiency. Whether or not this law reflects some underlying physical property of radio planets in general is not known;

indeed it was not our intention to suggest that such a property exists. However, the potential predictive value of the proposed relationship is, we feel, of some interest and will be tested further by Voyager spacecraft observations of Uranus in the coming year.

M. D. DESCH
M. L. KAISER

Goddard Space Flight Center,
Greenbelt,
Maryland 20771, USA

- Desch, M. D. & Kaiser, M. L. *Nature* **310**, 755-757 (1984).
- Voots, G. R., Gurnett, D. A. & Akasofu, S.-I. *J. geophys. Res.* **82**, 2259-2266 (1977).
- Gallagher, D. L. & D'Angelo, N. *Geophys. Res. Lett.* **8**, 1087-1089 (1981).
- Desch, M. D. & Rucker, H. O. *J. geophys. Res.* **88**, 8999-9006 (1983).
- Desch, M. D. & Barrow, C. H. *J. geophys. Res.* **89**, 6819-6823 (1984).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.